2-2010

Rare Decays in BaBar: Search for Decays of the Neutral B Meson into Two Leptons and Measurement of CP Asymmetry in Inclusive Radiative Decays of the B Meson into a Strange-quark System

Emmanuele Salvati
University of Massachusetts Amherst, emmanuele.salvati@gmail.com

RARE DECAYS IN *BABAR*: SEARCH FOR $B^0 \to \ell^+\ell'^-$ DECAYS AND MEASUREMENT OF CP ASYMMETRY IN INCLUSIVE $B \to X_s\gamma$ DECAYS.

A Dissertation Presented

by

EMMANUELE SALVATI

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

February 2010

Physics

RARE DECAYS IN *BABAR*: SEARCH FOR $B^0 \to \ell^+\ell'^-$ DECAYS AND MEASUREMENT OF CP ASYMMETRY IN INCLUSIVE $B \to X_s\gamma$ DECAYS.

A Dissertation Presented

by

EMMANUELE SALVATI

Approved as to style and content by:

Carlo Dallapiccola, Chair

Stephane Willocq, Member

Lorenzo Sorbo, Member

Grant Wilson, Member

Donald Candela, Department Chair
Physics

To my brother Carlo,
who taught me how to drive.

ABSTRACT

RARE DECAYS IN *BABAR*: SEARCH FOR $B^0 \to \ell^+\ell'^-$ DECAYS AND MEASUREMENT OF *CP* ASYMMETRY IN INCLUSIVE $B \to X_s\gamma$ DECAYS.

FEBRUARY 2010

EMMANUELE SALVATI

B.Sc., UNIVERSITÀ DI ROMA, LA SAPIENZA

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Carlo Dallapiccola

The purpose of this thesis is the study of rare decays with the *BABAR* detector. In the first part we present the study of the leptonic decays $B^0 \to e^+e^-$, $B^0 \to \mu^+\mu^-$ and $B^0 \to e^\pm\mu^\mp$, based on a dataset corresponding to $(383.6 \pm 4.2) \times 10^6$ $B\bar{B}$ pairs. We do not find evidence of any of the three decay modes, and obtain upper limit on the branching fractions, at 90% confidence level, of $\mathcal{BR}(B^0 \to e^+e^-) < 11.3 \times 10^{-8}$, $\mathcal{BR}(B^0 \to \mu^+\mu^-) < 5.2 \times 10^{-8}$, and $\mathcal{BR}(B^0 \to e^\pm\mu^\mp) < 9.2 \times 10^{-8}$.

In the second part we present the measurement of *CP* asymmetry in rare inclusive $B \to X_s\gamma$ decays on the recoil of fully-reconstructed hadronic B decays, using a dataset corresponding to $(465.0 \pm 5.1) \times 10^6$ $B\bar{B}$ pairs. We measure a value of $A_{CP} = -0.12 \pm 0.11 \pm 0.03$.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

INTRODUCTION

After ten years of running, the two B factories PEP-II and KEKB have collected an amount of data much larger than expected. The two main goals of the B physics experiments have been reached: the measurement of $\sin 2\beta$ [1] through the study of time dependent CP asymmetry in $b \to c\bar{c}s$ decays and the measurement of direct CP violation in $B \to K^+\pi^-$ decays [2]. The consistency with the Standard Model (SM) is now evident.

Nevertheless, the description of the world provided by the SM is incomplete. It does not explain the flavor properties of the particles, like their masses or the number of families, it does not justify accidental symmetries as the conservation of lepton and baryon numbers. Moreover, there are potential quadratic corrections to the Higgs mass which would make it difficult to explain its value in the SM.

Flavor Changing Neutral Current (FCNC) processes are not yet strongly constrained, so there is still room for New Physics (NP) effects. They represent the ideal framework for direct NP searches. However, they are highly suppressed with respect to the dominant tree decays of the B meson. Because of that, a large amount of data is needed to have a significant discriminating power.

Waiting for the next generation of B experiments (*i.e.* LHCb at CERN or, possibly, a Super B factory in Italy), we can use the available experimental information to reduce the allowed space of parameters for possible NP scenarios.

The study of $B_{d,s} \to \ell\ell$ decays offers the possibility to probe NP in a complementary way with respect to $b \to s$ time dependent analyses [3]. In the latter case, time dependent CP parameters are sensitive to new weak phases originating from complex couplings between squarks of different flavors [4]. This implies that this class

of decays has no sensitivity to physics beyond the SM in a particular class of model, Minimal Flavor Violation [5] (MFV), in which NP is flavor blind. In this case, indeed, the CKM matrix is the only source of flavor mixing and NP only shows up as shifts in $B^0 - \bar{B}^0$ mixing and decay rates of rare decays. $B_{d,s} \to \ell\ell$ decays are among the first ones to be studied in MFV scenarios, as they can be calculated with relatively small hadronic corrections.

Direct *CP* asymmetry in $b \to s\gamma$ decays is a powerful tool in detection of eventual NP, because of possible effects that could appear directly in lowest order loops. The corresponding rates can be calculated reliably at quark level, as non-perturbative effects do not play a dominant role: they are, indeed, under control using heavy quark expansion and quark-hadron duality. On the other hand, inclusive $B \to X_s\gamma$ decays are a challenge for experiments: they require the separation of few signal events from a very big background. The technique adopted in this thesis consists of studying high energy photons recoiling against fully reconstructed B mesons, which significantly reduces the background. The drawback of this technique is that it requires very high statistics, due to the very small signal efficiency, which can be achieved only at B factories.

The two experimental analyses presented in this work rely on data collected with the *BABAR* detector, operating at the PEP-II collider at the SLAC National Accelerator Laboratory.

CHAPTER 1

LEPTONIC AND RADIATIVE B DECAYS IN THE STANDARD MODEL

1.1 The Standard Model

According to the Standard Model of particle physics, matter is constituted by fractional spin particles called *fermions*, interacting through the exchange of integer spin particles named *gauge bosons* [6].

Two kinds of interactions, strong and electroweak, are described by the SM. Strongly interacting fermions are called *quarks*, while fermions experiencing only the electroweak interactions are called *leptons*. The electroweak interaction manifests itself at low energy as two distinct forces, weak and electromagnetic.

Until now, 6 leptons and 6 quarks have been discovered, while the strong, weak and electromagnetic interactions have, respectively, 8 gauge bosons (gluons), 3 gauge bosons ($W^{\pm}$ and Z) and 1 gauge boson (photon). Properties of all known elementary particles are listed in Table 1.1.

Fermions are organized as multiplets of the three groups $SU(3)$, $SU(2)$ and $U(1)$, in such a way that the interactions can be described in terms of a *lagrangian density* $\mathcal{L}$, satisfying the $SU(3)\times SU(2)\times U(1)$ symmetry. Quantum chromodynamics (QCD) for strong interactions emerges from the $SU(3)$ symmetry, while the $SU(2)\times U(1)$ symmetry generates the electroweak interactions.

For a multiplet Ψ of fermions fields, the transformation under a group $\mathcal{G}$ (*gauge transformation*) is defined by:

$$\Psi \mapsto G\Psi\,, \tag{1.1}$$

Particle	Symbol	Mass	charge	spin
LEPTONS				
electron	e	0.511	-1	1/2
neutrino e	ν_e	~ 0	0	1/2
muon	μ	106	-1	1/2
neutrino μ	ν_μ	~ 0	0	1/2
tau	τ	1777	-1	1/2
neutrino τ	ν_τ	~ 0	0	1/2
QUARKS				
down	d	$3.5 \div 6.0$	-1/3	1/2
up	u	$1.5 \div 3.3$	2/3	1/2
strange	s	105	-1/3	1/2
charm	c	1270	2/3	1/2
bottom	b	4200	-1/3	1/2
top	t	171.3×10^3	2/3	1/2
GAUGE BOSONS				
W	$W^\pm$	80×10^3	± 1	1
Z	Z	91×10^3	0	1
photon	γ	0	0	1
gluons	g	0	0	1

Table 1.1. Known elementary particles and their properties: approximate mass in MeV, charge in units of the elementary charge, spin. For each fermion, an antifermion exists with opposite charge.

where $G \in \mathcal{G}$ is usually written in the form $G = e^{\frac{i}{2}\alpha_i(x)O^i}$, with O^i a set of operators called *generators* of $\mathcal{G}$.

The simplest transformation we can define is $\psi \mapsto e^{i\alpha}\psi$, where ψ is a fermion field and α is a constant (this is a transformation belonging to the global $U(1)$ group). The lagrangian density [1]:

$$\mathcal{L} = \bar{\psi} i\gamma^\mu \partial_\mu \psi \tag{1.2}$$

is invariant under this transformation. When switching to a local $U(1)$ symmetry, i.e. when introducing a spatial dependence of the parameter α, the lagrangian can be kept invariant by replacing the derivative ∂_μ by a covariant derivative D_μ defined by:

$$D_\mu \equiv \partial_\mu - igA_\mu(x)\,, \tag{1.3}$$

where g is a constant (coupling constant) and A_μ is a new field (gauge field), with a suitable gauge transformation, corresponding to a gauge boson. The formalism can be extended to more complicated groups like $SU(2)$ and $SU(3)$, by introducing the corresponding generators in the transformations and organizing the fermions in multiplets of the gauge groups.

We are mostly interested in the electroweak sector. In order to reproduce the observed parity violation of weak interactions, the $SU(2) \times U(1)$ lagrangian is built by classifying the fermions in left handed doublets and right handed singlets of $SU(2)$. Each multiplet defines a *flavor family*:

$$\Psi_L \in \left\{ \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, \begin{pmatrix} u' \\ d' \end{pmatrix}_L, \begin{pmatrix} c' \\ s' \end{pmatrix}_L, \begin{pmatrix} t' \\ b' \end{pmatrix}_L \right\} \tag{1.4}$$

$$\Psi_R \in \{e_R, \mu_R, \tau_R, q_R\} \quad q \in \{u', d', c', s', t', b'\}\,. \tag{1.5}$$

[1] In this chapter we adopt the units where $\hbar = c = 1$.

Here, we indicate the quarks with a primed symbol since the electroweak eigenstates do not necessarily coincide with the mass eigenstates of Table 1.1.

Unfortunately, when introducing a mass term in the lagrangian, the $SU(2)\times U(1)$ symmetry is broken. Since the gauge symmetry ensures the physical predictions of the model to be finite [7], the SM does not allow an explicit symmetry breaking but introduces a *spontaneous symmetry breaking*. A new doublet of scalar fields (Higgs doublet) is introduced and the corresponding invariant lagrangian is written by introducing a potential that is symmetric under $SU(2)\times U(1)$, but whose minimum (corresponding to the vacuum state) has to be chosen among infinite equivalent non-symmetric states. As a result, at low energy, the Higgs field can be replaced in the lagrangian with its vacuum expectation value (v.e.v) plus perturbations around it: the v.e.v. breaks the symmetry and produces mass terms for the gauge bosons, while perturbations (identified with a Higgs boson) can produce additional interactions.

The 3 gauge bosons arising from the $SU(2)$ symmetry and the gauge boson arising from $U(1)$ are replaced by the mass eigenstates $W^{\pm}$, Z (massive) and photon (massless), where the Z and the photon are a mix of the $U(1)$ boson and one $SU(2)$ boson. The fermion-fermion-boson interactions can be written as:

$$\begin{aligned}\mathcal{L}_{int} &= \frac{g}{\sqrt{2}}\left[\overline{u}'_i\gamma^\mu\frac{1-\gamma^5}{2}d'_i+\overline{\nu}_i\gamma^\mu\frac{1-\gamma^5}{2}e_i\right]W^+_\mu+h.c.+\\ &- eJ^{em}_\mu A^\mu-\frac{g}{2\cos\theta_W}J^Z_\mu Z^\mu,\end{aligned} \tag{1.6}$$

where we defined the neutral currents:

$$J^{em}_\mu = \sum_f q_f\overline{f}\gamma_\mu f\,, \tag{1.7}$$

$$J^Z_\mu = \sum_f \overline{f}\gamma_\mu(g^f_V-g^f_A\gamma^5)f\,, \tag{1.8}$$

with

$$g_V^f = I_f^3 - 2q_f \sin^2\theta_W \,, \tag{1.9}$$

$$g_A^f = I_f^3 \,. \tag{1.10}$$

In these equations, $u_i' \in \{u', c', t'\}$ (up sector), $d_i' \in \{d', s', b'\}$ (down sector), $e_i \in \{e, \mu, \tau\}$, $\nu_i \in \{\nu_e, \nu_\mu, \nu_\tau\}$; f is a generic fermion with electric charge q_f and I_f^3 is the third component of the weak isospin ($+1/2$ for u_i and ν_i, $-1/2$ for d_i and e_i); θ_W is the Weinberg angle defining the mixing that produces the Z and the photon. Notice that right handed neutrinos would not interact with anything in this model, so are not introduced in Eq. (1.6).

Yukawa couplings between the Higgs and the fermion doublets are also introduced in order to produce fermion mass terms from the Higgs vacuum. These couplings can mix the flavor families, so that, as already mentioned, the flavor and mass eigenstates can be different. The diagonalization of the mass matrices requires unitary transformations that can be written as:

$$\begin{pmatrix} u' \\ c' \\ t' \end{pmatrix} = U_U \begin{pmatrix} u \\ c \\ t \end{pmatrix}, \qquad \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = U_D \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.11}$$

When introducing these transformations in the couplings with the charged W, we get terms like:

$$\begin{pmatrix} \bar{u}' & \bar{c}' & \bar{t}' \end{pmatrix} \gamma^\mu(1-\gamma^5) \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} \bar{u} & \bar{c} & \bar{t} \end{pmatrix} \gamma^\mu(1-\gamma^5) U_U^\dagger U_D \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.12}$$

The unitary matrix $V_{CKM} = U_U^\dagger U_D$ in general is not diagonal and is known as the *Cabibbo-Kobayashi-Maskawa* (CKM) matrix [8]:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4)\,, \tag{1.13}$$

We have used the Wolfenstein parameterization [9], which is an approximate parameterization where each element is expanded as a power series in the small parameter $\lambda = |V_{us}| = 0.22$. This matrix introduces couplings among quarks of different families. Notice that the same operation for the neutral current interactions gives, for the up sector:

$$\begin{pmatrix} \bar{u}' & \bar{c}' & \bar{t}' \end{pmatrix} \gamma^\mu(1-\gamma^5) \begin{pmatrix} u' \\ c' \\ t' \end{pmatrix} = \begin{pmatrix} \bar{u} & \bar{c} & \bar{t} \end{pmatrix} \gamma^\mu(1-\gamma^5) U_U^\dagger U_U \begin{pmatrix} u \\ c \\ t \end{pmatrix}. \tag{1.14}$$

and, given the unitarity of U_U, it follows $U_U^\dagger U_U = \mathbf{1}$ and there is no mixing. The same holds obviously for the down sector. FCNC are then forbidden at tree level in the SM. They are allowed at loop level thanks to the different quark masses that avoid exact cancellations of the contributions involving different quarks in the loops.

1.1.1 A few notes on renormalization

Divergent integrals appear when performing calculations with loops. In order to remove these divergences from the theory, they are absorbed in the definition of the lagrangian fields and parameters (masses and couplings), with a technique called *renormalization*. In other words, the lagrangian depends on the unobservable *bare* parameters, while the final results depend only on the *renormalized* parameters, the only ones that can be actually measured (for a mass, for instance, we can write the redefinition in the form $m_0 = Z_m m$, being m_0 the bare mass and m the renormalized one). Since the divergences are absorbed in the definition of the renormalized param-

eters, they do not appear in the final result. This procedure introduces, in the masses and in the couplings, a dependence (called *running*) on the mass scale μ of the process taken into account. In particular, in the case of QCD, the coupling α_s becomes greater than 1 below a scale $\Lambda_{QCD} \sim O(1\ \text{GeV})$, precluding the possibility of performing perturbative calculations in this region. Another feature of QCD renormalization is the appearance of terms of the form $\alpha_s^n(\mu_0)(\log(\mu_0/\mu))^m$ $(m = n, n-1, \ldots)$, when running the couplings and the masses from a high scale μ_0 to a lower scale μ. The large logarithms $\log(\mu_0/\mu)$ compensate the smallness of $\alpha_s^n(\mu_0)$, thus the impact of higher order terms in any perturbation expansion becomes important. In order to obtain reliable results, the largest logarithms (up to $m = n$, up to $m = n-1$ and so on, according to the required precision of the calculation) need to be summed for all n. It can be done by means of the *renormalization group* (RG) *equations*. Thanks to these equations, the quantities measured at the scale μ_0 can be run down to the scale μ in such a way that large logarithms are automatically summed for all orders. For a mass $m(\mu)$ the equation can be written in the form:

$$\frac{dm(\mu)}{d\log\mu} = -\gamma_m(g(\mu))m(\mu)\,, \tag{1.15}$$

where $g(\mu)$ is the coupling and $\gamma_m(g(\mu))$ is called *anomalous dimension*:

$$\gamma_m(g(\mu)) = \frac{1}{Z_m}\frac{dZ_m}{d\log\mu}\,. \tag{1.16}$$

Such large logarithms – for example $\log(M_W/\mu)$ – arise when calculating QCD corrections to the electroweak processes, and also in this case RG equations can be written and allow to sum them for all orders.

1.2 Effective lagrangian formalism

Flavor physics is governed by the interplay of strong and weak interactions. One of the main difficulties in examining the observables in flavor physics is the influence of the strong interactions. For matrix elements dominated by long-distance strong interactions, there is no adequate quantitative estimate available in quantum field theory because hadronic corrections become non-perturbative. The resulting hadronic uncertainties restrict the opportunities in flavor physics significantly, in particular within the indirect search for new physics.

Problems involving multiple energy scales can be often simplified by the introduction of an effective field theory (EFT). For instance, when treating interactions mediated by a massive gauge boson much heavier than the scale at which the interaction is happening, it is possible to remove from the theory the degrees of freedom associated with the massive boson.

The general idea is that, at a scale lower than a cut-off Λ, only the low frequency components of a field can appear in the initial and final state, hence we do not need explicitly the presence of the high frequency fields in the lagrangian. Formally, one says that the corresponding degrees of freedom can be "integrated out" of the action integral [10].

The result is that an effective lagrangian can be introduced, depending only on the low frequency modes and expanded in terms of all the possible operators satisfying the symmetries of the original lagrangian:

$$\mathcal{L}_{\Lambda}^{eff} = \sum_i c_i Q_i \,. \tag{1.17}$$

where the coefficients c_i are called *Wilson coefficients* and are related to the couplings of the full theory. The operators Q_i can be classified according to their dimension. Considering that the dimension of the lagrangian has to be energy to the fourth power, we can write:

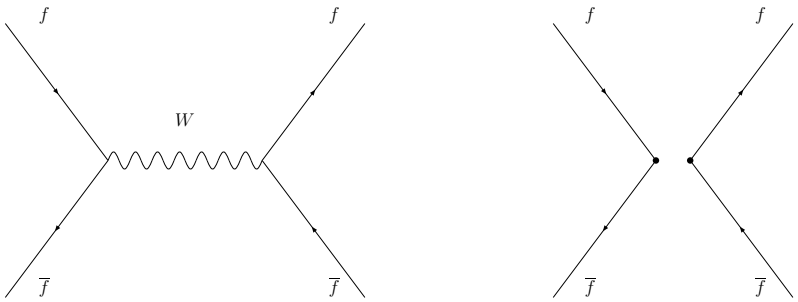

Figure 1.1. Feynman diagrams for the four-fermion weak interaction in the full theory (left) and in the Fermi theory (right).

$$\mathcal{L}_{\Lambda}^{eff} = \sum_{i} \frac{C_i}{\Lambda^{d_i-4}} Q_i \,. \tag{1.18}$$

where C_i is a dimensionless coefficient and d_i is the dimension of Q_i. From "naturalness" arguments, we can say that $C_i \sim O(1)$, unless some symmetry forces it to be much smaller or larger. Hence, just few operators with the lower dimensions are important in the lagrangian expansion.

In this context, assuming that there is some NP at some high scale, the SM itself can be treated as an EFT, with additional operators and corrections to the Wilson coefficients introduced by the NP. Hence, it is interesting to write the SM in the formalism of the EFTs. In doing that, considering that flavor physics is mainly interested in weak interactions at the GeV scale, we can also integrate out the degrees of freedom associated with the gauge boson field, by setting a cut-off $\Lambda = M_W$. This is the spirit of the Fermi theory of weak interactions, expressed in the modern language of the EFTs. The typical result is that the boson exchange diagram in the left of Figure 1.1 is replaced by the effective local four-fermion interaction in the right of the same figure.

Since each fermion line has dimension 3/2 in energy, the operators involved in this case are at least of order 6. So, we expect a M_W^{-2} scaling law for the Wilson coefficients at the leading order. Hence, it is useful to write the effective lagrangian as:

$$\mathcal{L}^{eff}_{weak} = -\frac{G_F}{\sqrt{2}} \sum C_i Q_i \,, \tag{1.19}$$

where $G_F/\sqrt{2} \equiv g^2/8M_W^2$ and the coefficients C_i are dimensionless for operators with dimension 6.

At tree level the effective lagrangian for the four-fermion process reads:

$$\mathcal{L}^{eff}_{4f} = -\frac{G_F}{\sqrt{2}} J^-_\mu J^{+\mu} \,, \tag{1.20}$$

where:

$$J^+_\mu \;=\; \sum_{ij} V_{ij}\, \bar{u}_i \gamma_\mu (1-\gamma_5) d_j + \sum_i \bar{\nu}_i \gamma_\mu (1-\gamma_5) e_i \,, \qquad J^-_\mu = (J^+_\mu)^\dagger \,. \tag{1.21}$$

The Wilson coefficients are computed by means of a matching procedure: the amplitude for simple processes are calculated in both the full and effective theories and the results are compared. Since the Wilson coefficients are process independent, the ones obtained by matching a specific process can then be used in the calculations for the others. This procedure can be performed including QCD corrections up to the desired order $O(\alpha^n)$. In this case, new operators arise and their color structure becomes important. Moreover, a QCD renormalization scale μ has to be introduced, of the same order of magnitude of the momenta involved in the process under study. The scale enters the matrix elements of Q_i and such a dependence is compensated by an analogous dependence in the Wilson coefficients.

Now an effective hamiltonian can be built. In the SM, the operators to be used can be classified in 6 categories. In the notation of [11]:

Current-Current Operators:

$$Q_1 = (\bar{s}_i c_j)_{\mathrm{V-A}} (\bar{c}_j b_i)_{\mathrm{V-A}} \qquad\qquad Q_2 = (\bar{s}_i c_i)_{\mathrm{V-A}} (\bar{c}_j b_j)_{\mathrm{V-A}}$$

QCD-Penguins Operators:

$$Q_3 = (\bar{s}_i b_i)_{\mathrm{V-A}} \sum_q (\bar{q}_j q_j)_{\mathrm{V-A}} \qquad Q_4 = (\bar{s}_i b_j)_{\mathrm{V-A}} \sum_q (\bar{q}_j q_i)_{\mathrm{V-A}}$$
$$Q_5 = (\bar{s}_i b_i)_{\mathrm{V-A}} \sum_q (\bar{q}_j q_j)_{\mathrm{V+A}} \qquad Q_6 = (\bar{s}_i b_j)_{\mathrm{V-A}} \sum_q (\bar{q}_j q_i)_{\mathrm{V+A}}$$

Electroweak-Penguins Operators:

$$Q_7 = \frac{3}{2} (\bar{s}_i b_i)_{\mathrm{V-A}} \sum_q e_q (\bar{q}_j q_j)_{\mathrm{V+A}} \qquad Q_8 = \frac{3}{2} (\bar{s}_i b_j)_{\mathrm{V-A}} \sum_q e_q (\bar{q}_j q_i)_{\mathrm{V+A}}$$
$$Q_9 = \frac{3}{2} (\bar{s}_i b_i)_{\mathrm{V-A}} \sum_q e_q (\bar{q}_j q_j)_{\mathrm{V-A}} \qquad Q_{10} = \frac{3}{2} (\bar{s}_i b_j)_{\mathrm{V-A}} \sum_q e_q (\bar{q}_j q_i)_{\mathrm{V-A}}$$

Magnetic and Chromo-magnetic Penguins Operators:

$$Q_{7\gamma} = \frac{\alpha}{2\pi} m_b \bar{s}_i \sigma^{\mu\nu} (1+\gamma_5) b_i F_{\mu\nu} \qquad Q_{8G} = \frac{\alpha_s}{2\pi} m_b \bar{s}_i \sigma^{\mu\nu} (1+\gamma_5) T^a_{ij} b_j G^a_{\mu\nu}$$

$\Delta S = 2$ **and** $\Delta B = 2$ **Operators:**

$$Q(\Delta S = 2) = (\bar{s}_i d_i)_{V-A} (\bar{s}_j d_j)_{V-A} \qquad Q(\Delta B = 2) = (\bar{b}_i d_i)_{V-A} (\bar{b}_j d_j)_{V-A}$$

Semi-Leptonic Operators:

$$Q_{7V} = (\bar{s}_i d_i)_{V-A} (\bar{e}e)_V \qquad Q_{7A} = (\bar{s}_i d_i)_{V-A} (\bar{e}e)_A$$
$$Q_{9V} = (\bar{b}_i s_i)_{V-A} (\bar{e}e)_V \qquad Q_{10A} = (\bar{b}_i s_i)_{V-A} (\bar{e}e)_A$$
$$Q^\nu_L = (\bar{s}_i b_i)_{V-A} (\bar{\nu}\nu)_{V-A} \qquad Q^\mu_L = (\bar{s}_i b_i)_{V-A} (\bar{\mu}\mu)_{V-A}$$

The examples provided above illustrate only the structure of the operators, while the quark content can change according to the process under study. Here, $(\bar{f}f)_{V\pm A} =$

$\bar{f}\gamma^{\mu}(1 \pm \gamma^5)f$ (indices for quarks indicate the color), m_b is the b quark mass, $F_{\mu\nu}$ and $G_{\mu\nu}$ are the electromagnetic and the gluon field strength tensors, respectively, and T^a are the Gell-Mann matrices.

The effective hamiltonian arising from this set of operators is usually written by making the impact of the quark mixing explicit, through the CKM matrix elements:

$$\mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}} \sum_i V^i_{CKM} C_i(\mu) Q_i \,, \tag{1.22}$$

where the sum runs over all the operators and V^i_{CKM} is the suitable CKM factor (e.g. $V^*_{ts}V_{tb}$ for $b \to s$ transitions). The amplitude for a process will be given by [11]:

$$\begin{aligned} \mathcal{A}(M \to F) &= \langle F|\mathcal{H}_{eff}|M\rangle = \frac{G_F}{\sqrt{2}} \sum_i V^i_{CKM} C_i(\mu) \langle F|Q_i(\mu)|M\rangle & (1.23) \\ &= \sum_i B_i \eta^i_{QCD} V^i_{CKM} F_i & (1.24) \end{aligned}$$

The last step comes from a rearrangement of the operators, where the *master functions* F_i's [12] are linear combinations of the Wilson coefficients at a high scale $\mu_0 \sim O(M_w, m_t)$, the B_i's parameterize the corresponding rearrangement of the matrix elements (and often include non-perturbative effects) and the η^i_{QCD}'s contain factors summarizing the evolution of the coefficients from the high to the low scale. Each of the F_i's (also known as *Inami-Lim functions*) represents a precise kind of box or penguin diagram [13], making the expressions for the decay amplitudes more intuitive. For instance, we will have Inami-Lim functions associated with the Z^0-penguin diagrams (i.e. electroweak and semileptonic penguin diagrams mediated by a Z^0 boson), the γ-penguin diagrams (as before, but mediated by a photon), magnetic penguin diagrams and so on.

The Inami-Lim functions can depend in general on a set υ of several parameters, but in the SM the dominant contributions depend only on $x_t = m_t^2/M_W^2$ [2]. Hence, we will write $F_i(\upsilon)$ when talking about the general form of these functions and just $F_i(x_t)$ when referring to their SM expressions.

From a practical point of view, the effective hamiltonian can be calculated as follows. At first, one defines the Inami-Lim functions at the high scale μ_0 and, with a matching procedure, translates them into the Wilson coefficients at the same scale. Since $\mu_0 \sim O(M_W) \gg \Lambda_{QCD}$, QCD perturbation theory can be used. At this point, it is possible to write the RG equations for the Wilson coefficients (which assume a form similar to Eq. (1.15)), and use them to evolve the coefficients from μ_0 to μ, with large logarithms automatically summed. Finally, non-perturbative methods have to be used to evaluate the matrix elements at the low scale. Notice these three features of this method:

1. the possibility of working in a well perturbative regime at M_W;
2. the possibility of resumming large logarithms in the running;
3. the difficulty of including the scale dependence in the non-perturbative calculations, so that the μ dependence of the Wilson coefficients is not completely compensated and the final results can show an unphysical μ dependence.

As already anticipated, with this formalism the inclusion of NP effects in the theory is easier. They can appear as:

- corrections to the master functions;
- new operators not present in the SM;

[2]Here, m_t is intended as the renormalized top mass at the μ_0 scale.

- new complex phases in the master functions;
- new sources of mixing not controlled by the CKM matrix.

Hence, the comparison between experiments and SM predictions can be done by setting some constraint on:

- corrections to the master functions or the Wilson coefficients;
- Wilson coefficients of new operators;
- *CP* violation and mixing effects not expected in the SM.

The absence of deviations from the SM in measurements of *CP* violation and mixing processes forced the theoreticians to concentrate on models where no new operators, no new phases and no new sources of mixing are present. This scenario, where all NP effects can be parameterized through corrections to the master functions (or equivalently to the SM Wilson coefficients), is called Minimal Flavor Violation (MFV) [5].

1.3 Leptonic B Decays

The SM does not allow FCNC at tree-level, and decays of this kind are predicted to have very small branching fractions. This makes rare decays particularly interesting for the detection of possible new physics (NP) beyond the SM, such as supersymmetry [14] (SUSY): loop contributions from heavy partners of the SM particles predicted in these models might induce, for certain decay modes, branching fractions significantly larger than the values predicted by the SM.

The leptonic decays $B^0 \to \ell^+\ell'^-$ (where $\ell^+\ell'^-$ stands for e^+e^-, $\mu^+\mu^-$ or $e^\pm\mu^\mp$; charge conjugation is implied throughout) are particularly interesting among rare decays, since a prediction of the decay rate in the context of the SM can be obtained with a relatively small error, due to the limited impact of long-distance hadronic

corrections [15]. In the SM, $B^0 \to \ell^+\ell^-$ decays proceed through diagrams such as those shown in Figure 1.2. These contributions are highly suppressed since they involve a $b \to d$ transition and require an internal quark annihilation within the B meson. The decays are also helicity suppressed by factors of $(m_\ell/m_B)^2$, where m_ℓ is the mass of the lepton and m_B the mass of the B meson.

In addition, B^0 decays to leptons of two different flavors violate lepton flavor conservation, so they are forbidden in the SM. This feature provides a handle to discriminate among different NP models [16].

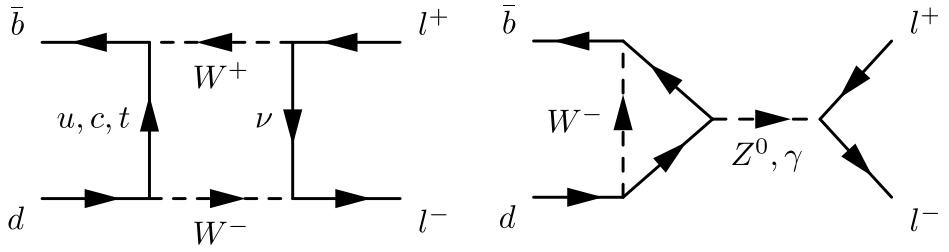

Figure 1.2. Representative Feynman diagrams for $B^0 \to \ell^+\ell^-$ in the Standard Model.

1.3.1 Effective Hamiltonian

In the SM there is only one operator of dimension six with a non-zero contribution to $B^0 \to \ell^+\ell^-$. It is:

$$Q = (\bar{b}d)_{V-A} \cdot (\bar{\ell}\ell)_{V-A} = (\bar{b}_L \gamma^\mu d_L) \cdot (\bar{\ell}\gamma_\mu\gamma_5\ell) \tag{1.25}$$

From this starting point, the effective hamiltonian can be written as [17]:

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} \frac{\alpha}{2\pi \sin^2\Theta_W} \left[\lambda_c F(x_c) + \lambda_t F(x_t)\right]. \tag{1.26}$$

The parameters λ_i are products of CKM matrix elements: $\lambda_t = V_{tb}^* V_{td}$, $\lambda_c = V_{cb}^* V_{cd}$. Values of these parameters are reported in Table 1.2.

	$B_s^0 \to \ell^+\ell^-$	$B^0 \to \ell^+\ell^-$
λ_c	$\sim \lambda^2$	$\sim \lambda^3$
λ_t	$\sim \lambda^2$	$\sim \lambda^3$

Table 1.2. CKM matrix parameters λ_c and λ_t, expressed in terms of the Wolfenstein parameter λ.

The parameters x_i are ratios of quark masses over the W-boson mass: $x_i = m_i^2/M_W^2$; the functions $F(x_i)$ are the master functions for this decay, which are monotonically increasing functions of x_i. They describe the dependence on the masses m_i of the internal quarks of the loop, which can be *up*, *charm* or *top* – with the most important contribution coming from the top – and contain the QCD corrections. The GIM (Glashow, Iliopoulos, Maiani) [18] suppression implies that for small values $x_i \ll 1$, the $F(x_i)$ are quadratic functions of the masses. We can make the approximation:

$$F(x) \sim x \ln x, \tag{1.27}$$

which has the following consequences:

- we can neglect the contribution from the up quark, for which it is $F(x_u) \sim 0$. We can state that the scales involved in this process are of the order of the top or charm quark mass, which are much bigger than the QCD scale (Λ_{QCD}). This is typical of short-range interactions;

- the relative importance of the charm and top quarks is given by the corresponding CKM matrix elements and by:

$$\frac{F(x_c)}{F(x_t)} \approx O(10^{-3}) \ll 1. \tag{1.28}$$

Given that λ_c and λ_t are of the same order of magnitude (see Table 1.2), we can neglect the contribution of the charm quark. With the above considerations, we can re-write the effective hamiltonian in the following way:

$$\mathcal{H}_{\text{eff}} = -\frac{G_F}{\sqrt{2}}\frac{\alpha}{2\pi \sin^2 \Theta_W} V_{tb}^* V_{td} \cdot Y(x_t) \cdot Q + \text{h.c.} \tag{1.29}$$

where:

- h.c. stands for hermitian conjugate;
- Q is the dimension-six operator defined in Eq. (1.25);
- $Y(x_t)$ takes into account the QCD correction [11]:

$$Y(x) = Y_0(x) + \frac{\alpha_s}{4\pi} Y_1(x); \tag{1.30}$$

- replacing the d quark with the s quark we get the corresponding $B_s^0 \rightarrow \ \ell^+\ell^-$.

The decay $B_s^0 \rightarrow \ \ell^+\ell^-$ has higher branching fraction than the corresponding $B^0 \rightarrow \ell^+\ell^-$ because of the parameters λ_c and λ_t, as can be seen from Table 1.2.

1.3.2 Branching fraction of $B^0 \rightarrow \ \ell^+\ell^-$

The branching fraction for $B_s^0 \rightarrow \ \ell^+\ell^-$ (B^0 is the same with the replacement of the quark s with the quark d) decays can be written starting from the effective hamiltonian of Eq. (1.29):

$$\mathcal{B}(B_s^0 \rightarrow \ \ell^+\ell^-) = \tau(B_s)\frac{G_F^2}{\pi}\left(\frac{\alpha}{4\pi \sin^2 \Theta_W}\right)^2 F_{B_s}^2 m_\ell^2 m_{B_s}^2 \sqrt{1 - 4\frac{m_\ell^2}{m_{B_s}^2}} |V_{tb}^* V_{ts}|^2 Y^2(x_t). \tag{1.31}$$

The QCD correction is included in the term $Y^2(x_t) = Y_0^2(x_t) + \frac{\alpha}{2\pi} Y_0(x_t) Y_1(x_t)$; the B-meson decay constant is defined as $\langle 0 \mid (\bar{b}s)_{V-A,\mu} \mid B_s(p) \rangle = iF_B p_\mu$; $\tau(B_S)$ is the mean lifetime of the B_s meson. If we use the following values:

- $\alpha = 1/129$;
- $\sin^2 \Theta_W = 0.23$;
- $Y(x) = \eta_Y Y_0(x)$, with $\eta_Y = 1.026 \pm 0.006$ being the QCD correction and $Y_0(x_t) = 0.784\ x_t^{0.76}$,

we get, for $B_s^0 \to\ \mu^+\mu^-$:

$$\mathcal{B}(B_s^0 \to\ \mu^+\mu^-) = 3.1 \cdot 10^{-9} \left[\frac{\tau(B_s)}{1.6\ \text{ps}}\right] \left[\frac{F_{B_s}}{0.21\ \ \text{GeV}}\right]^2 \left[\frac{|V_{ts}|}{0.040}\right]^2 \left[\frac{m_t(\mu_t)}{166\ \ \text{GeV}}\right]^{3.12}. \quad (1.32)$$

Taking the central values for $\tau(B_s)$, F_{B_s}, $|V_{ts}|$ and $m_t(m_t)$, assuming the value for the CKM ratio of $|V_{td}/V_{ts}| = (4.0 \pm 0.8) \cdot 10^{-2}$ and varying μ_t in the interval $100 \leq \mu_t \leq 300\ \ \text{GeV}$, we obtain:

$$6.8 \cdot 10^{-11} \leq \mathcal{B}(B^0 \to\ \mu^+\mu^-) \leq 3.4 \cdot 10^{-10}. \quad (1.33)$$

The $B^0 \to \ell^+\ell'^-$ decays are sensitive to NP in a large set of models with MFV, in which the NP lagrangian is flavor blind at the typical mass scale of new heavy states, with reduced effects on flavor physics at the B mass scale [19]. In the context of MFV models, NP corrections to $B^0 \to \ell^+\ell'^-$ are characterized by interesting correlations with other rare decays for a particular choice of some fundamental parameters (as in the case of small [20] or large [21] $\tan\beta$ in SUSY models with MFV). A precise determination of the decay rate of $B^0 \to \ell^+\ell'^-$ would allow different NP scenarios to be disentangled.

As shown in Table 1.3, the present experimental limits on $B^0 \to \ell^+\ell'^-$ are several orders of magnitude larger than SM expectations. Nevertheless, improved experimental bounds will restrict the allowed parameter space of several NP models.

Decay mode	$B^0 \to e^+e^-$	$B^0 \to \mu^+\mu^-$	$B^0 \to e^\pm\mu^\mp$
SM prediction	1.9×10^{-15}	8.0×10^{-11}	0
BABAR [22]	6.1×10^{-8}	8.3×10^{-8}	18×10^{-8}
Belle [23]	1.9×10^{-7}	1.6×10^{-7}	1.7×10^{-7}
CDF [24]	-	2.3×10^{-8}	-
CLEO [25]	8.3×10^{-7}	6.1×10^{-7}	15×10^{-7}

Table 1.3. The expected branching fractions in the Standard Model [15] and the available upper limits (UL) at 90% C.L.

1.4 Radiative B Decays

Since the b quark mass is much larger than the typical scale of the strong interaction Λ_{QCD}, long-distance strong interactions could be generally taken under control, thanks to the expansion in that heavy mass [26]. A good example of this approach is inclusive rare B decays, since they are theoretically clean and represent a laboratory of perturbative QCD.

In particular, the decay width $\Gamma(B \to X_s\gamma)$ is well approximated by the partonic decay rate $\Gamma(b \to s\gamma)$, which can be analyzed in RG-improved perturbation theory:

$$\Gamma(B \to X_s\gamma) = \Gamma(b \to s\gamma) + \Delta^{nonpert.}. \tag{1.34}$$

Non-perturbative effects, $\Delta^{nonpert.}$, play a subdominant role and are under control thanks to the heavy mass expansion and the assumption of quark–hadron duality [27](see Sec. 1.4.1).

The SM leading order diagrams for $b \to s\gamma$, shown in Figure 1.3, are called *penguin diagrams*.

In contrast to the exclusive rare B decay modes, the inclusive ones are theoretically clean observables, because no specific model is needed to describe the hadronic final states.

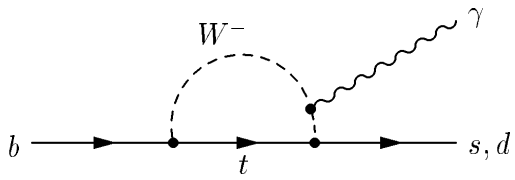

Figure 1.3. Feynman diagram for the electromagnetic penguins $b \to s\gamma$ and $b \to d\gamma$. The photon can be emitted from the W (shown) or from any of the quarks.

The inclusive modes $B \to X_s\gamma$ can be measured in electron–positron colliders (B factories, CLEO) because of their kinematic constraints and their controlled backgrounds, while they are more difficult to measure at hadronic machines.

1.4.1 Heavy quark effective theory

The heavy-quark effective theory (HQET) is constructed to provide a simplified description of processes where a heavy quark interacts with light degrees of freedom predominantly by the exchange of soft gluons. In these systems typical momenta exchanged between the heavy and light constituents are of order $\Lambda_{\rm QCD} \sim 0.2$ GeV. The heavy quark is surrounded by a complicated, strongly interacting cloud of light quarks, antiquarks, and gluons. In this case it is the fact that the Compton wavelength of the heavy quark $\lambda_Q \sim 1/m_Q$ is much smaller than the size of the hadron $R_{\rm had} \sim 1/\Lambda_{\rm QCD}$, which leads to simplifications. Resolving the quantum numbers of the heavy quark would require a hard probe, the soft gluons exchanged between the heavy quark and the light constituents can only resolve distances much larger than λ_Q. Therefore, the light degrees of freedom are blind to the flavor – hence, to the mass – and spin orientation of the heavy quark. They experience only its color field, which extends over large distances because of confinement, whereas relativistic effects such as color magnetism vanish as $m_Q \to \infty$.

These hadronic bound states are therefore characterized by a large separation of mass scales. The goal of the HQET is to separate the physics associated with these two scales, in such a way that all dependence on the heavy-quark mass becomes

explicit. The framework in which to perform this separation is the operator product expansion (OPE) [28, 29] (see Section 1.2).

After the separation of short- and long-distance phenomena a big portion of the relevant physics (i.e. all short-distance effects) could be computed using perturbation theory and RG techniques, taking under control all logarithmic dependence on the heavy-quark mass, and it may happen that the long-distance physics simplifies due to approximate symmetries, which imply non-trivial relations between observables.

Compared with most effective field theories, in which the degrees of freedom of a heavy particle are removed completely from the low-energy theory, the HQET is special in that its purpose is to describe the properties and decays of hadrons which do contain a heavy quark. Hence, it is not possible to remove the heavy quark completely from the effective theory. What is possible is to integrate out the "small components" in the full heavy-quark spinor, which describe the fluctuations around the mass shell.

The ordinary QCD lagrangian for a heavy-quark field Ψ with mass m

$$\mathcal{L} = \bar{\Psi} i \not{D} \Psi - m \bar{\Psi} \Psi, \tag{1.35}$$

with the covariant derivative

$$D_\mu = \partial_\mu - igT^a A^a_\mu, \tag{1.36}$$

could be expressed [30, 31] in terms of the large- and small-component fields, h_v and H_v,

$$\Psi(x) = e^{-imv\cdot x} \left(h_v(x) + H_v(x) \right), \tag{1.37}$$

where the heavy-quark momentum has been decomposed as

$$p = mv + k, \tag{1.38}$$

with v being the 4-velocity of the heavy *hadron*. After selecting the large kinematical part of the momentum, mv, the remaining component k is determined by the soft QCD bound state interactions, and thus $k = \mathcal{O}(\Lambda_{QCD}) \ll m$.

Eq. 1.35 could be re-written as:

$$\mathcal{L}_{\text{eff}} = \bar{h}_v \, iv\cdot D \, h_v + \frac{1}{2m_Q} \, \bar{h}_v \, (iD_\perp)^2 \, h_v + \frac{g_s}{4m_Q} \, \bar{h}_v \, \sigma_{\mu\nu} \, G^{\mu\nu} \, h_v + O(1/m_Q^2) \, . \qquad (1.39)$$

The first term in Eq. 1.39 describes the "residual" QCD dynamics of the heavy quark, after the kinematic dependence on m is separated out. Since there is no longer any reference to the mass m, the only parameter to distinguish quark flavors, this term is flavor symmetric: the dynamics is the same for b and c quarks in the static limit. Since the operator $v \cdot D$ contains no γ-matrices, which would act on the spin degrees of freedom, the leading HQET lagrangian also exhibits a spin symmetry. This corresponds to the decoupling of the heavy-quark spin in the $m \to \infty$ limit [32].

The second term describes the non-relativistic kinetic energy arising from the off-shell residual motion of the heavy quark, and the third represents the chromo-magnetic coupling of the heavy-quark spin to the gluon field.

The following definitions are used in literature [31]:

$$\lambda_1 \equiv \frac{\langle B|\bar{h}(iD)^2 h|B\rangle}{2m_B} \qquad \lambda_2 \equiv \frac{1}{6}\frac{\langle B|\bar{h}g\sigma \cdot G h|B\rangle}{2m_B}. \qquad (1.40)$$

Inclusive decay rates determine the probability for the decay of a particle into the sum of all possible final states with a given set of global quantum numbers. From a theoretical point of view, inclusive decays of hadrons containing a heavy quark offer two advantages [33, 34]. First, bound-state effects related to the initial state (such as the "Fermi motion" of the heavy quark inside the hadron [35, 36]) can be accounted for in a systematic way using the heavy-quark expansion. Secondly since

the energy released into the final state by the decay of the heavy b quark is large compared to the QCD scale, the final hadronic state needs not be dominated by a few sharp resonances. If resonances are indeed unimportant, then there is a factorization between the short distance part of the decay (the disappearance of the b quark) and the long distance part (the eventual hadronization of the decay products). This factorization implies that for sufficiently inclusive quantities it is enough to consider the short distance part of the process, with the subsequent hadronization taking place with unit probability. This factorization, known as *local parton-hadron duality*, is an example of a crucial assumption which lies outside of the HQE itself. Local duality must hold as $m_b \to \infty$ with all other masses held fixed. In this limit, wavelengths associated with the b quark decay are arbitrarily short and cannot interfere coherently with the hadronization process. On the other hand, it is not known how to estimate the size of corrections to local duality for m_b large but finite. There is no analog of the heavy quark expansion appropriate to this question, and no way to estimate systematically deviations from the limit $m_b \to \infty$. Although an expansion in powers $1/m_b$ in the calculation of inclusive quantities is incorporated, the behavior of this expansion does not address directly the issue of violations of duality. The duality hypothesis, while entirely reasonable for inclusive B decays, is not independently verifiable except by the direct confrontation of theoretical calculations with the data.

The inclusive decay width of a heavy hadron Γ_H, could be expressed [31], using the optical theorem, as:

$$\Gamma_H = \frac{1}{2m_H}\langle H|\mathcal{T}|H\rangle \equiv \langle \mathcal{T}\rangle, \tag{1.41}$$

where the transition operator $\mathcal{T}$ is defined as

$$\mathcal{T} = \text{Im } i\int d^4x\, T\,\mathcal{H}_{\text{eff}}(x)\mathcal{H}_{\text{eff}}(0), \tag{1.42}$$

with $\mathcal{H}_{\text{eff}}$ being the effective weak Hamiltonian. Eqs. (1.41) and (1.42) express the total decay rate as the absorptive part of the forward scattering amplitude $H \to H$

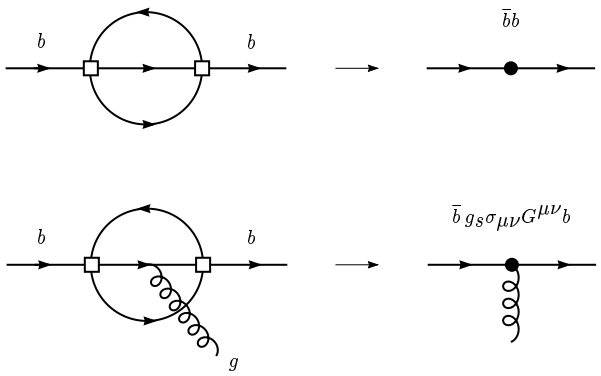

Figure 1.4. Perturbative contributions to the transition operator **T** (left), and the corresponding operators in the operator product expansion (right). The open squares represent a four-fermion interaction of the effective weak lagrangian $\mathcal{L}_{\text{eff}}$, while the black circles represent local operators in the $1/m_b$ expansion.

under the action of $\mathcal{H}_{\text{eff}}$. They could be re-written in a more directly understandable form by inserting a complete set of states $\sum_\alpha |X\rangle\langle X|$ between the two factors of $\mathcal{H}_{\text{eff}}$ in Eq. 1.42 and removing the T-product by explicitly taking the absorptive part. This yields

$$\Gamma_H \sim \sum_\alpha \langle H|\mathcal{H}_{\text{eff}}|X\rangle\langle X|\mathcal{H}_{\text{eff}}|H\rangle, \tag{1.43}$$

where one immediately recognizes the decay rate as the modulus squared of the decay amplitude (summed over all final states X). The reason to introduce (1.42) is that the T-product, by means of Wick's theorem, allows for a direct evaluation in terms of Feynman diagrams.

In order to compute Γ_H an operator product expansion is applied to Eq. 1.42, resulting in a series of local operators of increasing dimension. The coefficients of these operators are correspondingly suppressed by increasing powers of $1/m_b$.

The leading contributions to the transition operator are shown in Figure 1.4. The result for $B \to X_s\gamma$ decays is:

$$\Gamma(\bar{B} \to X_f) = \frac{G_F^2 m_b^5}{192\pi^3} \left\{ c_3^f \left(1 + \frac{\lambda_1 + 3\lambda_2}{2m_b^2}\right) + c_5^f \frac{6\lambda_2}{m_b^2} + \ldots \right\}. \tag{1.44}$$

The main result of the HQE for inclusive decay rates is the observation that the free quark decay (*i.e.* the parton model) provides the first term in a systematic $1/m_b$ expansion [33]. For dimensional reasons, the corresponding rate is proportional to the fifth power of the b-quark mass.

The non-perturbative corrections, which arise from bound-state effects inside the B meson, are suppressed by at least two powers of the heavy-quark mass, *i.e.* they are of relative order $(\Lambda_{\rm QCD}/m_b)^2$. The absence of first-order power corrections is a consequence of the equations of motion, as there is no independent gauge-invariant operator of dimension four that could appear in the operator product expansion.

The fact that bound-state effects in inclusive decays are strongly suppressed explains *a posteriori* the success of the parton model in describing such processes [37, 38].

1.4.2 $B \to X_s\gamma$ theoretical predictions

The general Eq. 1.44 can describe the inclusive $B \to X_s\gamma$ decays. In this case it could be written as

$$\Gamma_{B\to X_s\gamma} = \frac{G_F^2 m_b^5}{32\pi^4} |V_{tb}V_{ts}|^2 C_7^2(m_b)(1 + \frac{1}{2}\lambda_1 - \frac{9}{2}\lambda_2). \tag{1.45}$$

where $C_7^2(m_b)$ is the Wilson coefficient for the dipole-operator shown in Figure 1.5. The resulting $\mathcal{O}(\Lambda_{QCD}^2/m_b^2)$ non perturbative correction amounts to around 3% with respect to the partonic decay [39].

The Wilson coefficient in Eq. 1.45 encodes information on the short-distance QCD effects due to hard gluon exchanges between the quark lines of the leading one-loop electroweak diagrams (Figure 1.6). Such effects enhance the branching ratio of $B \to X_s\gamma$ by roughly a factor of three [40, 41]. Moreover, it is sensitive to the top quark mass, and more generally, to any kind of new physics beyond the standard model.

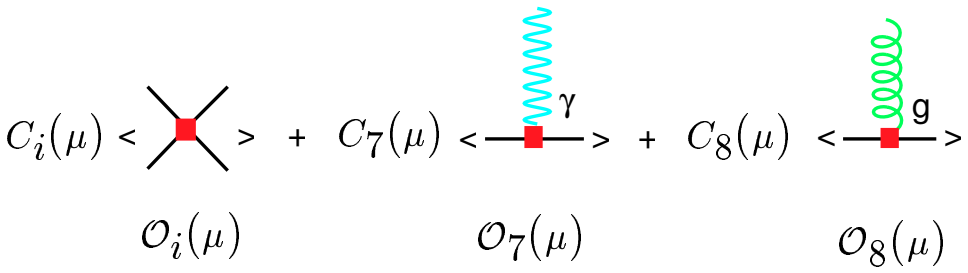

Figure 1.5. Effective Hamiltonian in the case of $B \to X_{s,d}\gamma$.

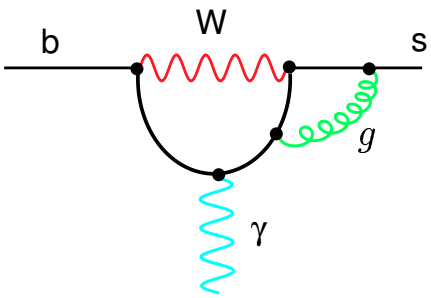

Figure 1.6. QCD corrections to the decay $b \to s\gamma$.

The leading order (LO) $\Gamma(b \to s\gamma)$ result [42] was dominated by a large renormalization scale dependence of about $\pm 25\%$. A next to leading order calculation (NLO) was needed, which was completed in 1997, thanks to the effort of many different groups ([43, 44, 45]). The theoretical error of the previous LO result was substantially reduced to $\pm 10\%$ and the central value of the partonic decay rate increased by about 20%.

Including the QED and the non-perturbative corrections discussed above, the theoretical prediction for the $B \to X_s\gamma$ branching ratio [46] is:

$$\mathcal{B}(B \to X_s\gamma) = (3.32 \pm 0.30) \times 10^{-4}, \tag{1.46}$$

where the error has two sources: the uncertainty regarding the μ scale dependences and the uncertainty due to the input parameters. In the latter the uncertainty due to the parameter m_c/m_b is dominant. This prediction almost coincides with the prediction of Kagan and Neubert [47].

An extra enhancement of the branching ratio is shown in reference [48]. This is due to two factors. An enhancement of QCD logarithms is due to the b-quark mass evolution in the top-quark sector. Quark mass effects were further analyzed, in particular the definitions of the quark masses m_c and m_b in the two-loop matrix element of the four-quark operators $\mathcal{O}_{1,2}$. Since the charm quark in the matrix elements $\mathcal{O}_{1,2}$ is dominantly off-shell, it is argued that the running charm mass should be chosen instead of the pole mass. The latter choice was used in all previous analyses [43, 45, 46, 47]. Using the running quark masses is rather important and leads to a $+11\%$ shift of the central value of the $B \to X_s\gamma$ branching ratio.

With the new choice of the charm mass renormalization scheme, the theoretical prediction for the 'total' branching ratio is

$$\mathcal{B}(B \to X_s\gamma)_{E\gamma > 1.6\,\mathrm{GeV}} = (3.60 \pm 0.30) \times 10^{-4}, \tag{1.47}$$

1.5 $B \to X_s\gamma$ decay model

In order to simulate the $B \to X_s\gamma$ decay it is necessary to determine the shape of the photon and m_{X_s} spectrum. Theoretical predictions are based on a non-resonant model [47], *i.e.* the resonances in the m_{X_s} spectrum have widths exceeding their spacing and thus overlapping. These predictions must be understood in the sense of quark-hadron duality. In the high region of the hadronic X_s mass spectrum, a large number of final states are kinematically accessible, thus the photon and hadronic mass spectra are similar to those predicted by HQE. On the other hand, in the low-mass region, the invariant mass of the hadronic final state is of the order $\Lambda^2_{\rm QCD}$, implying that the photon energy is very close to the kinematic endpoint. In this case, there could be resonance structures due to low-lying kaon states, not predicted by HQE. This is the case of the $K^*(892)$, whose width has been measured in [49] to be 50.8 MeV. This resonance needs to be incorporated in the decay model with width and branching ratio set to the measured value. This issue will be addressed in the next two sections.

1.5.1 Non-resonant contribution

In the B rest frame, the non-resonant spectrum can be described equivalently in terms of the photon energy or the invariant mass of the hadronic system X_s. The relation between the two spectra can be related from kinematics:

$$E_\gamma = \frac{m_B^2 - m_{X_s}^2}{2m_B}. \tag{1.48}$$

It is necessary to recall that the theoretical predictions for the photon energy and hadronic mass spectra must be understood in the sense of quark–hadron duality. In particular, the true hadronic mass spectrum in the low-mass region may have resonance structures due to low-lying kaon states. Two kinematic regions could be defined: the "endpoint region" and the "resonance region".

State H	M_H [GeV/c^2]	Γ_H [MeV]	E_γ [GeV]
$K\,(n\pi)$	≥ 0.629	continuum	≤ 2.60
$K^*(892)$	0.894	50	2.56 ± 0.01
$K_1(1270)$	1.273	90	2.49 ± 0.02
$K_1(1400)$	1.402	174	2.45 ± 0.05
$K^*(1410)$	1.412	227	2.45 ± 0.06
$K_2^*(1430)$	1.428	103	2.45 ± 0.03
$K_2(1580)$	1.580	110	2.40 ± 0.03
$K_1(1650)$	1.650	150	2.38 ± 0.05
$K^*(1680)$	1.714	323	2.36 ± 0.10
$K_2(1770)$	1.773	186	2.34 ± 0.06

Table 1.4. Mean masses and widths of the lowest-lying hadronic states accessible in $B \to X_s\gamma$ decays, and the corresponding photon energies (errors refer to changing M_H by $\pm\Gamma_H$).

The endpoint region of the photon energy spectrum is characterized by the condition that $E_\gamma^{\rm max} - E_\gamma = O(\bar{\Lambda})$, where $\bar{\Lambda} = m_B - m_b$. It is in this region that the effects of Fermi motion [47] – the motion of the b quark inside the B meson – are relevant and determine the shape of the spectrum. In the endpoint region, the invariant mass of the hadronic final state is of order $m_B\bar{\Lambda} \gg \Lambda^2_{\rm QCD}$, implying that a large number of final states are kinematically accessible. Under such circumstances, local quark–hadron duality ensures that the photon and hadronic mass spectra are similar to the corresponding inclusive spectra predicted by the HQE.

However, in the resonance region the invariant mass of the hadronic final state is of order $\Lambda^2_{\rm QCD}$, implying that the photon energy is very close to the kinematic endpoint: $E_\gamma^{\rm max} - E_\gamma = O(\Lambda^2_{\rm QCD}/m_B)$. In this case, HQE does not allow model-independent predictions for the structure of the individual resonance contributions.

The X_s state can decay through a number of resonances given in Table 1.4. There are six resonances plus a continuum contribution feeding the photon spectrum in the energy interval between 2.4 and 2.6 GeV. Hence, an average over this interval should be calculable using global quark–hadron duality, although a much finer resolution cannot be obtained. In the hadronic mass spectrum, the $K^*(892)$ peak is clearly

separated from the rest; however, the next resonances already have widths exceeding the level spacing and hence are overlapping. Therefore only this resonance will be considered separately. The prescription given in [47] consists of a single Breit–Wigner peak for the $K^*(892)$ followed by a continuum above a threshold $M_{\rm cont}$, which is dual to the higher resonance contributions and given by the inclusive spectrum calculated using HQE.

This gives

$$\frac{\mathrm{d}\mathcal{B}}{\mathrm{d}M_H} = \frac{2M_H N_{K^*}\,\mathcal{B}(B\to K^*\gamma)}{(M_H^2 - m_{K^*}^2)^2 + m_{K^*}^2\Gamma_{K^*}^2} + \Theta(M_H - M_{\rm cont})\,\frac{\mathrm{d}\mathcal{B}_{\rm incl}}{\mathrm{d}M_H}\,, \tag{1.49}$$

where

$$N_{K^*} = \frac{m_{K^*}\Gamma_{K^*}}{\arctan\left(\dfrac{m_{K^*}}{\Gamma_{K^*}}\right) + \dfrac{\pi}{2}}. \tag{1.50}$$

is the normalization of the Breit–Wigner distribution.

The continuum threshold $M_{\rm cont}$ is then fixed by the requirement that the total branching ratio be the same as that predicted by HQE, yielding the condition

$$\int_0^{M_{\rm cont}} \mathrm{d}M_H\,\frac{\mathrm{d}\mathcal{B}_{\rm incl}}{\mathrm{d}M_H} = \mathcal{B}(B\to X_s\gamma)\Big|_{E_\gamma > E_{\rm cont}} = \mathcal{B}(B\to K^*\gamma)\,, \tag{1.51}$$

where $E_{\rm cont} = \frac{1}{2}(m_B^2 - M_{\rm cont}^2)/m_B$.

The ratio of resonant over non-resonant branching fractions, as well as the cut-off mass, depends on input parameters. The result is shown in Figure 1.7.

In this analysis, the $B\to K^*(892)\gamma$ over the non-resonant ratio is fixed to the experimental measurements, taken from the Particle Data Group (PDG) [50]. The weighted average of the *BABAR* measurements for the two charge states [49] gives, for the resonant branching ratio:

$$\mathcal{B}(B\to K^*\gamma) = (4.2\pm 0.6)\times 10^{-5}. \tag{1.52}$$

The world weighted average, for the non-resonant branching ratio, is:

$$\mathcal{B}(B \to X_s\gamma) = (3.34 \pm 0.38) \times 10^{-4}. \tag{1.53}$$

This gives a $B \to K^*(892)\gamma$ over the non-resonant ratio of $\sim 12\%$.

The parameters, from which the shape function is computed, see Eq. B.4, are set to the following values: $m_b = 4.65$ GeV/c^2, $\mu_\pi^2 = 0.3$ GeV^2, $m_B = 5.2788$ GeV/c^2. The minimum energy of the photon is given by $E_\gamma \geq \frac{1}{2}(1-\delta)m_b$ where the cutoff parameter δ is set to be 0.9. The renormalization scale μ_b is set to the b-quark mass m_b. The ratio of the charm and beauty quark masses $z = m_c/m_b$ is set to 0.22 in agreement with the most recent theoretical calculation [48].

The m_{X_s} cut-off is computed in agreement with the Kagan-Neubert prescription described above, resulting $m_{X_s}^{cut-off} = 1.15$ GeV/c^2. Making use of Eq. 1.48 , this leads to an upper cut-off for the energy photon at $E_\gamma = 2.51$ GeV.

In this analysis we rely on the signal model primarily to compute our efficiency and the effect on it varying the model; only the shape and the ratio of resonant over non-resonant matters for this.

This model is also used for the optimization of the selection criteria. This is the only place where an assumed inclusive branching fraction matters. But in this context, neither that assumption nor using a "wrong" model can cause any bias; at worst they result in slightly non-optimum selection cuts.

The error on the experimental measurement, and on the used model will be taken into account in the systematics effect.

1.5.2 Resonant contribution

The branching fractions of the decay $B^\pm \to K^{*\pm}(892)\gamma$ and $B^0 \to K^{*0}(892)\gamma$ have been measured [49] to be:

$$\mathcal{B}(B^\pm \to K^{*\pm}(892)\gamma) = (3.83 \pm 0.62(\text{stat.}) \pm 0.22(\text{syst.})) \times 10^{-5} \ ,$$

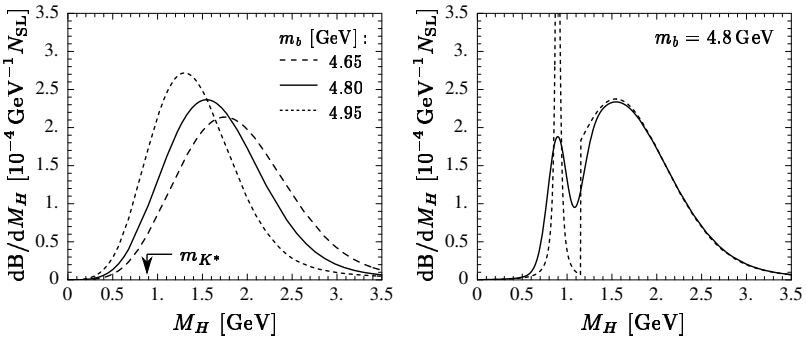

Figure 1.7. The m_{X_s} spectra taken from reference [47] ($m_{X_s} = m_H$). a) The spectra using Eq. 1.48 for different m_b choices; b) one of the m_{X_s} spectra modified to include the $K^*(892)$ resonance [47].

$$\mathcal{B}(B^0 \to K^{*0}(892)\gamma) = (4.23 \pm 0.40(\text{stat.}) \pm 0.22(\text{syst.})) \times 10^{-5} \ .$$

These numbers are not well predicted by theory because they require the difficult calculation of a heavy-to-light form factor at $q^2 = 0$. In fact recent calculations give values of $\sim 7 \times 10^{-5}$ which are large compared to the experimental measurement. As already discussed in the previous section the $B \to K^*\gamma$ is modeled by a Breit–Wigner function with peak and width fixed, from the experiments, to be respectively 892 MeV and 50.8 MeV and the branching ratio set to the weighted average of the *BABAR* measurements:

$$\mathcal{B}(B \to K^*\gamma) = (4.2 \pm 0.6) \times 10^{-5}. \tag{1.54}$$

1.6 CP asymmetry in $B \to X_s\gamma$ decays

The CKM mechanism that predicts *CP* violation introducing one single phase has passed its first precision test in the golden B mode, $B_d \to J/\psi K_S$, at the 5% level. Nevertheless, there is still room for non-standard *CP* phases, especially in the FCNC $\Delta S = 1$ modes.

The direct *normalized CP* asymmetries of the inclusive decay modes is given by [3]:

$$\alpha_{CP} = \frac{\Gamma(b \to s\gamma) - \Gamma(\bar{b} \to \bar{s}\gamma)}{\Gamma(b \to s\gamma) + \Gamma(\bar{b} \to \bar{s}\gamma)}.$$

Such an asymmetry can be different from zero only if the decay is due to two or more amplitudes with different strong and weak phases.

It is important to distinguish between $b \to s\gamma$ and $b \to d\gamma$ in making these measurements. The SM predicts in fact much larger *CP* asymmetries in $b \to d\gamma$ ($\approx 10\%$), but in the sum of $b \to s\gamma$ and $b \to d\gamma$ the *CP* asymmetries exactly cancel in the U-spin limit. Even if the U-spin limit does not make much sense at quark level – it requires the masses of the d quark and of the s quark to be equal – we can use the U-spin symmetry for hadronic matrix elements, in which case we get:

$$\Delta\Gamma(\bar{B} \to X_s\gamma) + \Delta\Gamma(\bar{B} \to X_d\gamma) = \Delta\Gamma_s + \Delta\Gamma_d = 0, \tag{1.55}$$

where $\Delta\Gamma_q = \Delta\Gamma(\bar{B} \to X_q\gamma) = \Gamma(\bar{B} \to X_q\gamma) - \Gamma(B \to X_{\bar{q}}\gamma)$. U-spin breaking effects can be accounted for, yielding to:

$$\Delta\Gamma(\bar{B} \to X_s\gamma) + \Delta\Gamma(\bar{B} \to X_d\gamma) = b_{\rm inc}\Delta_{\rm inc} \tag{1.56}$$

where the right-hand side is written as a product of a 'U-spin breaking' term $b_{\rm inc}$ and a 'typical size' $\Delta_{\rm inc}$ of the *CP* violating rate difference. A rough estimate of $b_{\rm inc}$ gives a value of the order $|b_{\rm inc}| \sim m_s^2/m_b^2 \sim 5 \times 10^{-4}$, leading to the estimate:

$$|\Delta\mathcal{B}(B \to X_s\gamma) + \Delta\mathcal{B}(B \to X_d\gamma)| \sim 1 \times 10^{-9}. \tag{1.57}$$

[3] This is the sign convention that is generally adopted in theory and experiment, which implies $\alpha_{CP}(B^0) = (\Gamma(\bar{B}^0 \to X_s^0\gamma) - \Gamma(B^0 \to X_{\bar{s}}^0\gamma))/(\Gamma(\bar{B}^0 \to X_s^0\gamma) + \Gamma(B^0 \to X_{\bar{s}}^0\gamma))$ and $(\alpha_{CP}(B^\pm) = (\Gamma(B^- \to X_s^-\gamma) - \Gamma(B^+ \to X_{\bar{s}}^+\gamma))/(\Gamma(B^- \to X_s^-\gamma) + \Gamma(B^+ \to X_{\bar{s}}^+\gamma))$.

This is also true in MFV extensions of the Standard Model. Theoretical NLL QCD predictions of the normalized CP asymmetries of the inclusive channels (see [51]) within the SM can be expressed by the approximate formula:

$$\begin{aligned} \alpha_{CP}(B \to X_s\gamma) &\approx 0.334 \times \Im[\epsilon_s] \approx +0.6\%\,, \\ \alpha_{CP}(B \to X_d\gamma) &\approx 0.334 \times \Im[\epsilon_d] \approx -16\%. \end{aligned} \tag{1.58}$$

where

$$\epsilon_s = \frac{V_{us}^* V_{ub}}{V_{ts}^* V_{tb}} \simeq -\lambda^2(\rho - i\eta), \qquad \epsilon_d = \frac{V_{ud}^* V_{ub}}{V_{td}^* V_{tb}} \simeq \frac{\rho - i\eta}{1 - \rho + i\eta}. \tag{1.59}$$

The two CP asymmetries are connected by the relative factor $\lambda^2\left((1-\rho)^2 + \eta^2\right)$. The small SM prediction for the *CP* asymmetry in the decay $B \to X_s\gamma$ is a result of three suppression factors. There is an α_s factor needed in order to have a strong phase; there is a CKM suppression of order λ^2 and there is a GIM suppression of order $(m_c/m_b)^2$ reflecting the fact that in the limit $m_c = m_u$ any *CP* asymmetry in the SM would vanish.

From an experimental point of view, there are several ways in which an α_{CP} analysis differs from the branching fraction analysis: the need to allow for mis-tagging; the possibility of asymmetries in the backgrounds or selection efficiencies; and the different way in which model-dependent uncertainties affect the result.

The flavor of the decaying $B \to X_s\gamma$ is determined from the tagging requirements on the non-signal B. A fraction ω of the tags for signal events will be assigned the wrong charge; ω is referred to as the *mistag fraction*. There are two contributions to ω. First, for B^0 and $\bar{B}^0$ decays there is a probability of an oscillation taking place before decay, leading to an incorrect flavor tag. The mistag fraction from this source is equal to the time-integrated B^0 - $\bar{B}^0$ mixing probability, χ, the world average value of which is 0.181 ± 0.004 [50]. This is an irreducible source of mis-tagging. Second, for both B^0 and $B^\pm$ mesons there are decays where the flavor of the b quark within the meson is incorrectly tagged.

In the analysis described in this thesis, one B is fully reconstructed in a hadronic decay, allowing the tagging of the flavor of the b quark with negligible mistag fraction and the separation of charged and neutral B mesons. This implies that the only contribution to the mistag ratio is due to neutral B oscillation. Also this measurement doesn't separate $B \to X_s\gamma$ decays from $B \to X_d\gamma$, although it would be possible with higher statistics by looking for the K_s in the X system.

The value of α_{CP}^{meas} , in neutral B mesons tag events, must be corrected for the mistag fraction to yield the underlying asymmetry, α_{CP}:

$$\alpha_{CP} = \frac{\alpha_{CP}^{\rm meas}}{(1-2\chi)} \, . \tag{1.60}$$

This means that the statistical precision of α_{CP} is diluted by a factor of $(1-2\chi)$. No correction needs to be applied in the case of charged B mesons tagged events.

The last consideration is the relative importance of statistical, systematic and model dependent errors in the α_{CP} analysis. As discussed in Section B the branching fraction analysis has a significant model dependence which decreases as more of the spectrum is included by lowering the minimum requirement on E_γ. For the α_{CP} measurement this model dependence cancels out to first order for any requirement on E_γ, therefore we can choose the cut on the photon energy by minimizing the corresponding experimental error on the CP asymmetry.

Since the SM rate is dominated by a single diagram, it predicts a CP asymmetry of less than 1%. The presence of new phases might lead to values of the asymmetry significantly different from zero [51, 52], as shown in Figure 1.8, without changing the inclusive branching fraction.

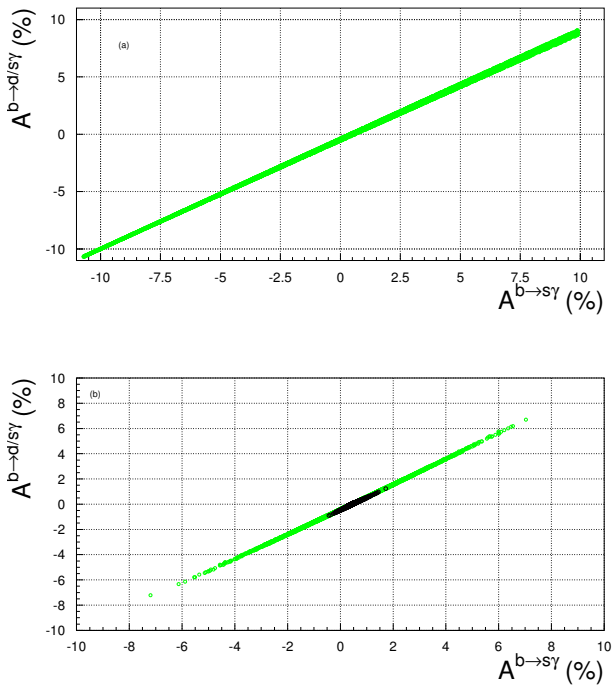

Figure 1.8. Correlation between A_{CP} in the untagged $B \to X_{s+d}\gamma$ and the tagged $B \to X_s\gamma$ modes. The top figure is obtained without the assumption of a particular model, whereas in the bottom one there is the extra requirement of minimal flavor violation (MFV). The dark shaded points are constrained by electic dipole moment (EDM).

CHAPTER 2

THE *BABAR* DETECTOR

The analysis described in this work has been performed on the data collected by the *BABAR* detector [53], operating at PEP-II the asymmetric B-Factory of the SLAC National Accelerator Laboratory. In this chapter we briefly describe PEP-II and the subsystems composing the *BABAR* detector.

2.1 Introduction

The PEP-II B-Factory is an e^+e^- asymmetric collider running at a center of mass energy of 10.58 GeV corresponding to the mass of the $\Upsilon(4S)$ resonance. The electron beam in the High Energy Ring (HER) has 9.0 GeV and the positron beam in the Low Energy Ring (LER) has 3.1 GeV. The $\Upsilon(4S)$ is therefore produced with a Lorentz boost of $\beta\gamma = 0.56$, allowing for a precise measurement of the time difference between the two B decays, as needed by the time-dependent CP violation measurements that are the primary goal of *BABAR*.

A longitudinal section of the *BABAR* detector is shown in Figure 2.1. The detector is composed by several subsystems. The tracking system, used to reconstruct the charged particles and the decay vertex, includes two different detectors: a Silicon Vertex Detector (SVT) and a Drift Chamber (DCH), both operating in a $1.5\,T$ magnetic field provided by a super-conducting solenoid. A detector of internal reflected Cherenkov light (DIRC) allows for particle identification (PID) and an Electromagnetic Calorimeter (EMC) is employed for photons reconstruction and energy mea-

surements. Finally, muon candidates are identified in the instrumented flux return (IFR) of the solenoid.

The detector is characterized by a hexagonal section and is divided in a central part (*barrel*) and two *end-caps*. The covered polar angle ranges from 350 mrad, in the forward, to 400 mrad in the backward directions (defined with respect to the high energy beam direction). The *BABAR* coordinate system has the z axis along the boost direction (or the beam direction): the y axis is vertical and the x axis is horizontal and goes toward the external part of the ring. In order to maximize the geometrical acceptance for $\Upsilon(4S)$ decays the whole detector is offset, with respect to the beam-beam interaction point (IP), by 0.37 m in the direction of the lower energy beam.

2.2 The PEP-II B-Factory

PEP-II is a system consisting of two accumulating asymmetric rings designed in order to operate at a center of mass energy of the $\Upsilon(4S)$ resonance mass, 10.58 GeV. Table 2.1 shows the design parameters and typical parameters in the 2007 run.

Parameters	Design	Typical
Energy HER/LER (GeV)	9.0/3.1	9.0/3.1
Current HER/LER (A)	0.75/2.15	1.96/3.03
# of bunch	1658	1732
bunch time separation (ns)	4.2	4.2
σ_{Lx} (μm)	110	157
σ_{Ly} (μm)	3.3	4.7
σ_{Lz} (μm)	9000	10000
Luminosity $(10^{33}$ cm$^{-2}s^{-1})$	3	12
Daily average integrated luminosity (pb^{-1}/d)	135	700

Table 2.1. PEP-II beam parameters. Design and typical values are quoted and are referred to the last year of machine running at the $\Upsilon(4S)$ resonance (2007).

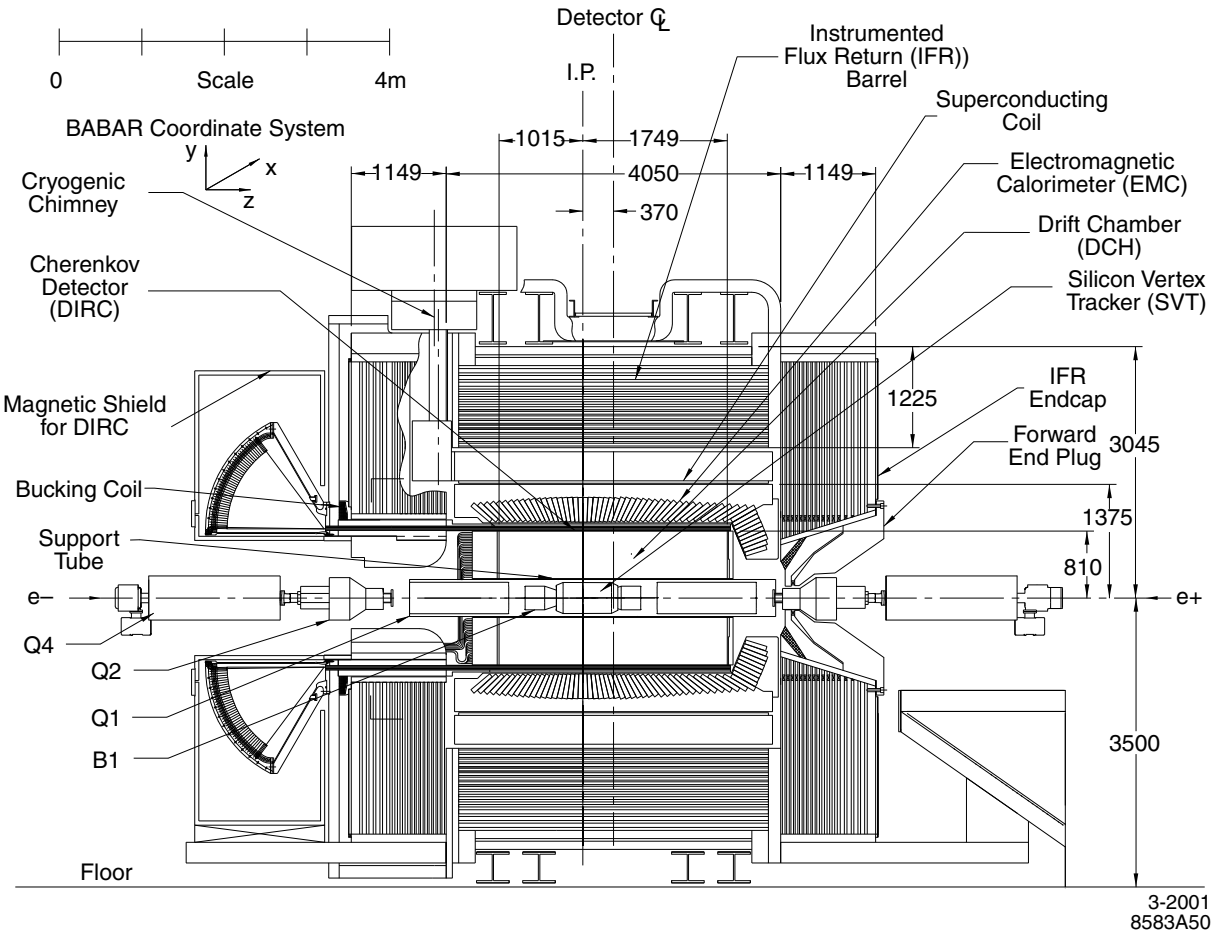

Figure 2.1. *BABAR* detector longitudinal section.

Data are mostly collected at $\Upsilon(4S)$ peak energy. Table 2.2 shows the active processes cross sections breakdown at peak energy. From now on the production of light quark pairs (u, d, s) and *charm* quark pairs will be referred to as *continuum production*. In order to study this non-resonant production $\sim 10\%$ of data is collected with a center of mass energy 40 MeV below the $\Upsilon(4S)$ mass value (*off-peak* data).

$e^+e^- \rightarrow$	Cross section (nb)
$b\bar{b}$	1.05
$c\bar{c}$	1.30
$s\bar{s}$	0.35
$u\bar{u}$	1.39
$d\bar{d}$	0.35
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	~ 40

Table 2.2. Various processes cross sections at $\sqrt{s} = M_{\Upsilon(4S)}$. *Bhabha* cross section is an effective cross section, within the experimental acceptance.

Recently, the machine has been run at different center of mass energies, corresponding to the $\Upsilon(2S)$ and $\Upsilon(3S)$ masses, in order to study rare bottomonium decays and look for unobserved bottomonium states [54]. An energy scan between 10.54 and 11.2 GeV has been also performed in order to look for exotic bottomonium states and assess the parameters of the $\Upsilon(5S)$ and $\Upsilon(6S)$ resonances [55].

The interaction region design, with the two beams crossing in a single interaction point with particles trajectories modified in order to have head on collisions, is realized with a magnetic field, produced by a dipole magnetic system, acting near the interaction point. The collision axis is off-set from the z-axis of the *BABAR* detector by about 20 mrad in the horizontal plane to minimize the perturbation of the beams by the solenoidal field. In this configuration the particles and the beams are kept far apart in the horizontal plane outside the interaction region and parassite collisions are minimized. Magnetic quadrupoles included inside the detector's magnetic field,

and hence realized in Samarium-Cobalt, are strongly focusing the beams inside the interaction region.

2.3 Tracking system

The charged particle tracking system is composed by a silicon vertex tracker (SVT) and a drift chamber (DCH): the main purpose of this tracking system is the efficient detection of charged particles and the measurement of their momenta and angles with high precision.

2.3.1 The silicon vertex tracker

The vertex detector has a radius of 20 cm from the primary interaction region: it is placed inside the support tube of the beam magnets and consists of five layers of double-sided silicon strip sensors detectors to provide five measurements of the positions of all charged particles with polar angles in the region $20.1° < \theta < 150°$. Because of the presence of a $1.5\,T$ magnetic field, the charged particle tracks with transverse momenta lower than $\sim 100\,\mathrm{MeV}/c$ cannot reach the drift chamber active volume. So the SVT has to provide stand-alone tracking for particles with transverse momentum less than $120\,\mathrm{MeV}/c$. For high momentum tracks, it also provides the measurement of track angles that is required to achieve design resolution for the Cherenkov angle in the DIRC.

In order to reach the required performances, the SVT is very close to the production vertex. It allows a very precise reconstruction of the charged particles trajectories on both longitudinal (z) and transverse directions. The longitudinal coordinate information is necessary to measure the decay vertex distance, while the transverse information allows a better separation between secondary vertices coming from decay cascades.

More precisely, the design of the SVT was carried out according to some important guidelines:

- The number of impact points of a single charged particle has to be greater than 3 to make a stand-alone tracking possible, and to provide an independent momentum measurement.
- The first three layers are placed as close as possible to the impact point to achieve the best resolution on the position along z of the B meson decay vertices.
- The two outer layers are close to each other, but comparatively far from the inner layers, to allow a good measurement of the track angles.
- The SVT must withstand 2 MRad of ionizing radiation: the expected radiation dose is 1 Rad/day in the horizontal plane immediately outside the beam pipe and 0.1 Rad/day on average.
- Since the vertex detector is inaccessible during normal detector operations, it has to be reliable and robust.

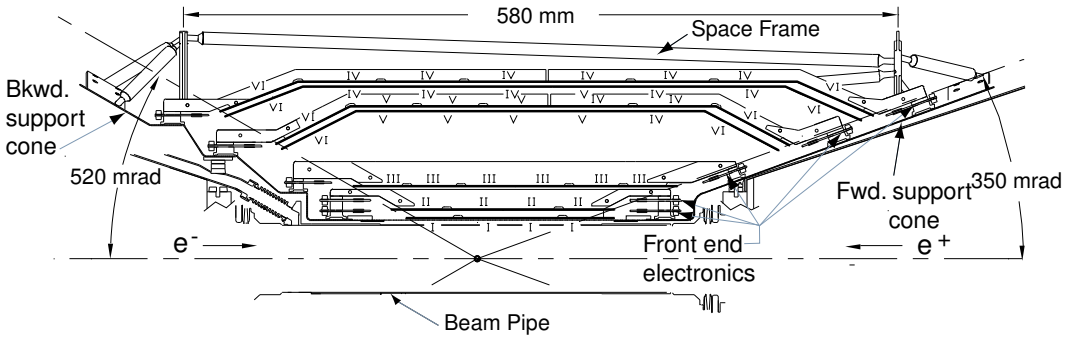

Figure 2.2. SVT schematic view: longitudinal section

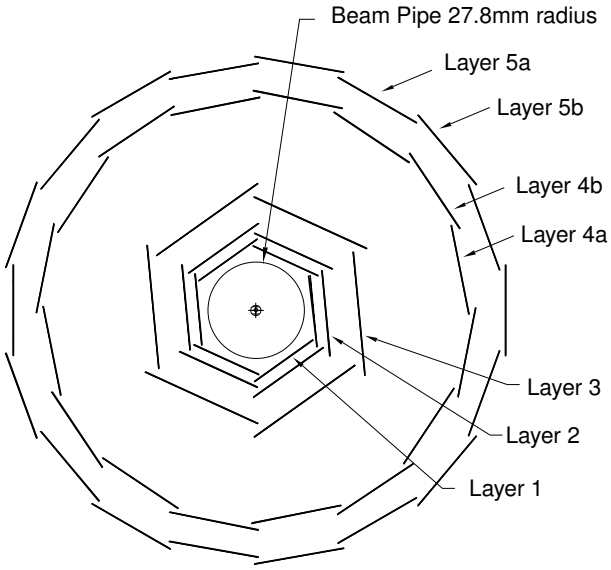

Figure 2.3. Cross-sectional view of the SVT in a plane perpendicular to the beam axis.

These guidelines have led to the choice of a SVT made of five layers of double-sided silicon strip sensors: the spatial resolution, for perpendicular tracks must be $10 - 15\,\mu m$ in the three inner layers and about $40\,\mu m$ in the two outer layers. The three inner layers perform the impact parameter measurement, while the outer layers are necessary for pattern recognition and low p_t tracking. The silicon detectors are double-sided (contain active strips on both sides) because this technology reduces the thickness of the materials the particles have to cross, thus reducing the energy loss and multiple scattering probability compared to single-sided detectors. The sensors are organized in modules (see Figure 2.2). The five layers of the SVT contain 340 silicon strip detectors with AC-coupled silicon strips.

Each detector is $300\,\mu m$-thick but sides range from $41\,mm$ to $71\,mm$ and there are 6 different detector types. Each of the three inner layers has a hexagonal transverse cross-section and it is made up of 6 detector modules, arrayed azimuthally around the beam pipe, while the outer two layers consist of 16 and 18 detector modules, respectively. The inner detector modules are barrel-style structures, while the outer

detector modules employ the novel arch structure in which the detectors are electrically connected across an angle. This arch design was chosen to minimize the amount of silicon required to cover the solid angle while increasing the solid angle for particles near the edges of acceptance: having incidence angles on the detector closer to 90 degrees at small dip angles insures a better resolution on impact points. The readout electronics is entirely mounted outside the active detector volume. The readout is organized in such a way that each module can be divided in a forward and a backward half-module, electrically separated.

The strips on the two sides of the rectangular detectors in the barrel regions are oriented parallel (ϕ strips) or perpendicular (z strips) to the beam line: in other words, the inner sides of the detectors have strips oriented perpendicular to the beam direction to measure the z coordinate (z-size), whereas the outer sides, with longitudinal strips, allow the ϕ-coordinate measurement (ϕ-side). In the forward and backward regions of the two outer layers, the angle between the strips on the two sides of the trapezoidal detectors is approximately 90° and the ϕ strips are tapered.

The inner modules are tilted in ϕ by 5°, allowing an overlap region between adjacent modules: this provides full azimuthal coverage and is convenient for alignment. The outer modules are not tilted, but are divided into sub-layers and placed at slightly different radii (see Figure 2.3).

The total silicon area in the SVT is $0.94\,m^2$ and the number of readout channels is about 150 000. The geometrical acceptance of SVT is 90% of the solid angle in the center of mass system and typically 80% is used in charged particle tracking.

The SVT efficiency can be calculated for each half-module by comparing the number of associated hits to the number of tracks crossing the active area of the half-module. The combined hardware and software efficiency is 97%.

The spatial resolution of SVT hits is calculated by measuring the distance (in the plane of the sensor) between the track trajectory and the hit, using high-momentum

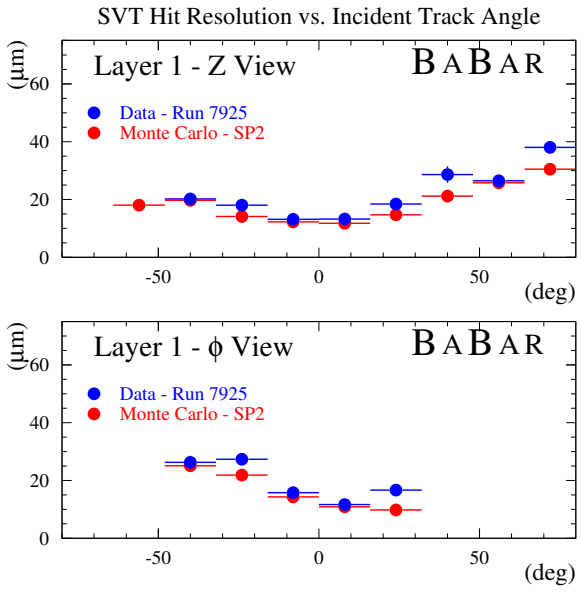

Figure 2.4. SVT hit resolution in the z and ϕ coordinate in microns, plotted as functions of the track incident angle in degrees.

tracks in two prong events: the uncertainty due to the track trajectory is subtracted from the width of the residual distribution to obtain the hit resolution. The track hit residuals are defined as the distance between track and hit, projected onto the wafer plane and along either the ϕ or z direction. The width of this residual distribution is then the SVT hit resolution. Figure 2.4 shows the SVT hit resolution for z and ϕ side hits as a function of the track incident angle: the measured resolutions are in very good agreement with the Monte Carlo expected ones. Over the whole SVT, resolutions range from $10 - 15\,\mu m$ (inner layers) to $30 - 40\,\mu m$ (outer layers) for normal tracks.

The double-sided sensors also provide up to ten measurements of dE/dx per track: with signals from at least four sensors, a 60% truncated mean dE/dx is calculated. For Minimum Ionizing Particles, the resolution on the truncated mean dE/dx is approximately 14%: a 2σ separation between kaons and pions can be achieved up to a momentum of 500 MeV/c and between kaons and protons beyond 1 GeV/c.

2.3.2 The drift chamber

The principal purpose of the DCH is the efficient detection of charged particles and the measurement of their momenta and angles with high precision. The DCH complements the measurements of the impact parameter and the directions of charged tracks provided by the SVT near the impact point (IP). At lower momenta, the DCH measurements dominate the errors on the extrapolation of charged tracks to the DIRC, EMC and IFR. The reconstruction of decay and interaction vertices outside of the SVT volume, for instance the K_S^0 decays, relies only on the DCH. For these reasons, the chamber should provide maximal solid angle coverage, good measurement of the transverse momenta and positions but also of the longitudinal positions of tracks with a resolution of $\sim 1\,mm$, efficient reconstruction of tracks at momenta as low as 100 MeV/c and it has to minimally degrade the performance of the calorimeter and particle identification devices (the most external detectors). For low momentum particles, the DCH is required to provide particle identification by measuring the ionization loss (dE/dx). A resolution of about 7% allows π/K separation up to 700 MeV/c. This PID measurement is complementary to that of the DIRC in the barrel region, while in the extreme backward and forward region, the DCH is the only device providing some discrimination of particles of different mass. The DCH should also be able to operate in presence of large beam-generated backgrounds having expected rates of about $5\,kHz$/cell in the innermost layers.

To meet the above requirements, the DCH is a 280 cm-long cylinder (see left plot in Figure 2.5), with an inner radius of 23.6 cm and an outer radius of 80.9 cm: it is bounded by the support tube at its inner radius and the particle identification device at its outer radius. The flat end-plates are made of aluminum: since the *BABAR* events will be boosted in the forward direction, the design of the detector is optimized to reduce the material in the forward end. The forward end-plate is thinner ($12\,mm$) in the acceptance region of the detector compared to the rear end-plate ($24\,mm$), and

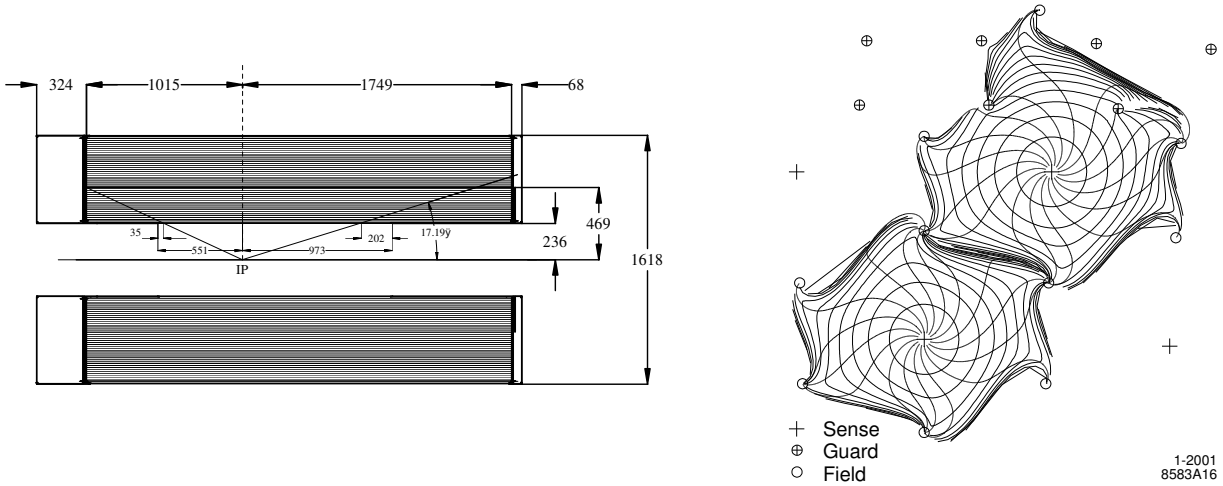

Figure 2.5. Side view of the *BABAR* drift chamber (the dimensions are in mm) and isochrones (i.e. contours of equal drift time of ions) in cells of layer 3 and 4 of an axial super-layer. The isochrones are spaced by $100\,ns$.

all the electronics is mounted on the rear end-plate. The device is asymmetrically located with respect to the IP: the forward length of 174.9 cm is chosen so that particles emitted at polar angles of 17.2° traverse at least half of the layers of the chamber before exiting through the front end-plate. In the backward direction, the length of 101.5 cm means that particles with polar angles down to 152.6° traverse at least half of the layers.

The inner cylinder is made of $1\,mm$ beryllium and the outer cylinder consists of two layers of carbon fiber glued on a Nomex core: the inner cylindrical wall is kept thin to facilitate the matching of SVT and DCH tracks, to improve the track resolution for high momentum tracks and to minimize the background from photon conversions and interactions. Material in the outer wall and in the forward direction is also minimized in order not to degrade the performance of the DIRC and the EMC.

The region between the two cylinders is filled up by a gas mixture consisting of Helium-isobutane (80% : 20%): the chosen mixture has a radiation length that is five times larger than commonly used argon-based gases. 40 layers of wires fill the DCH volume and form 7104 hexagonal cells with typical dimensions of 1.2×1.9 cm^2 along

the radial and azimuthal directions, respectively (see right plot in Figure 2.5). The hexagonal cell configuration has been chosen because approximate circular symmetry can be achieved over a large portion of the cell. Each cell consist of one sense wire surrounded by six field wires: the sense wires are $20\,\mu m$ gold-plated tungsten-rhenium, the field wires are $120\,\mu m$ and $80\,\mu m$ gold-plated aluminum. By using the low-mass aluminum field wires and the helium-based gas mixture, the multiple scattering inside the DCH is reduced to a minimum, representing less than $0.2\%X_0$ of material. The total thickness of the DCH at normal incidence is $1.08\%X_0$.

The drift cells are arranged in 10 super-layers of 4 cylindrical layers each. The super-layers contain wires oriented in the same direction: to measure the z coordinate, axial wire super-layers and super-layers with slightly rotated wires (*stereo*) are alternated. In the stereo super-layers a single wire corresponds to different ϕ angles and the z coordinate is determined by comparing the ϕ measurements from axial wires and the measurements from rotated wires. The stereo angles vary between ± 45 mrad and ± 76 mrad.

While the field wires are at ground potential, a positive high voltage is applied to the sense wires: an avalanche gain of approximately 5×10^4 is obtained at a typical operating voltage of $1960\,V$ and a $80:20$ helium:isobutane gas mixture.

In each cell, the track reconstruction is obtained by the electron time of flight: the precise relation between the measured drift time and drift distance is determined from sample of e^+e^- and $\mu^+\mu^-$ events. For each signal, the drift distance is estimated by computing the distance of closest approach between the track and the wire. To avoid bias, the fit does not include the hit of the wire under consideration. The estimated drift distances and the measured drift times are averaged over all wires in a layer.

The DCH expected position resolution is lower than $100\,\mu m$ in the transverse plane, while it is about $1\,mm$ in the z direction. The minimum reconstruction and momentum measurement threshold is about 100 MeV/c and it is limited by the DCH

inner radius. The design resolution on the single hit is about $140\,\mu m$ while the achieved weighted average resolution is about $125\,\mu m$. The top plot in Figure 2.6 shows the position resolution as a function of the drift distance, separately for the left and the right side of the sense wire. The resolution is taken from Gaussian fits to the distributions of residuals obtained from unbiased track fits: the results are based on multi-hadron events for data averaged over all cells in layer 18.

The specific energy loss (dE/dx) for charged particles through the DCH is derived from the measurement of the total charge collected in each drift cell: the specific energy loss per track is computed as a truncated mean from the lowest 80% of the individual dE/dx measurements. Various corrections are applied to remove sources of bias: these corrections include changes in gas pressure and temperature ($\pm 9\%$ in dE/dx), differences in cell geometry and charge collection ($\pm 8\%$), signal saturation due to space charge buildup ($\pm 11\%$), non-linearities in the most probable energy loss at large dip angles ($\pm 2.5\%$) and variation of cell charge collection as a function of the entrance angle ($\pm 2.5\%$).

The bottom plot in Figure 2.6 shows the distribution of the corrected dE/dx measurements as a function of track momenta: the superimposed Bethe-Bloch predictions have been determined from selected control samples of particles of different masses. The achieved dE/dx r.m.s. resolution for Bhabha events is typically 7.5%, limited by the number of samples and Landau fluctuations, and it is close to the expected resolution of 7%.

2.4 Cherenkov light detector

The Detector of Internally Reflected Cherenkov light is designed to provide a good K/π separation ($\sim 4\sigma$) in particular above $700\,\mathrm{MeV}/c$.

The particle identification in the DIRC is based on the Cherenkov radiation produced by charged particles crossing a material with a speed higher than light speed

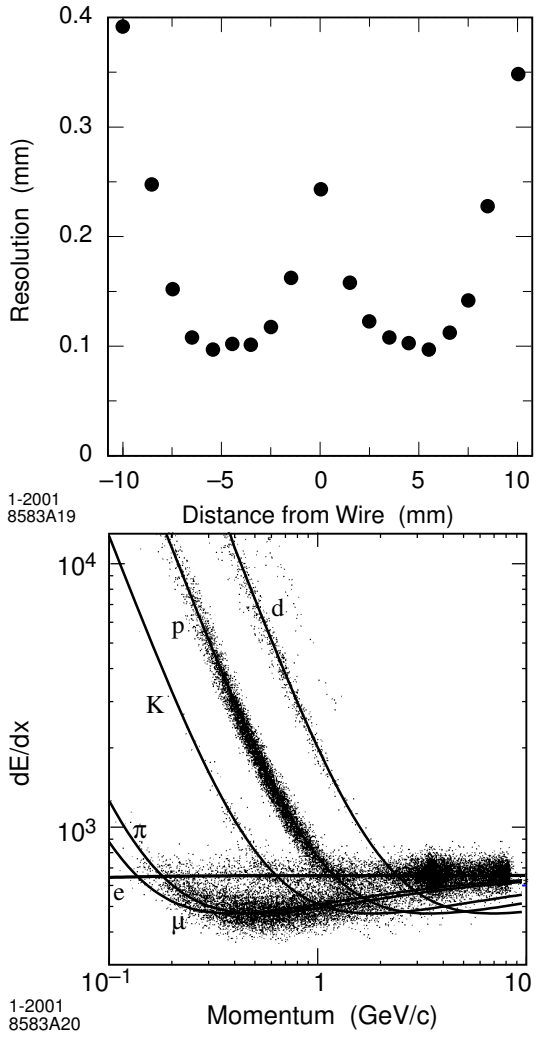

Figure 2.6. Top plot: DCH position resolution as a function of the drift chamber in layer 18, for tracks on the left and right side of the sense wire. The data are averaged over all cells in the layer. Bottom plot: measurement of dE/dx in the DCH as a function of the track momenta. The data include large samples of beam background triggers as evident from the high rate of protons. The curves show the Bethe-Bloch predictions derived from selected control samples of particles of different masses.

in that material. The angular opening of the Cherenkov radiation cone depends on the particle speed:

$$cos\theta_c = \frac{1}{n\beta} \tag{2.1}$$

where θ_c is the Cherenkov cone opening angle, n is the refractive index of the material and β is the particle velocity over c. The principle of the detection is based on the fact that the magnitudes of angles are maintained upon reflection from a flat surface.

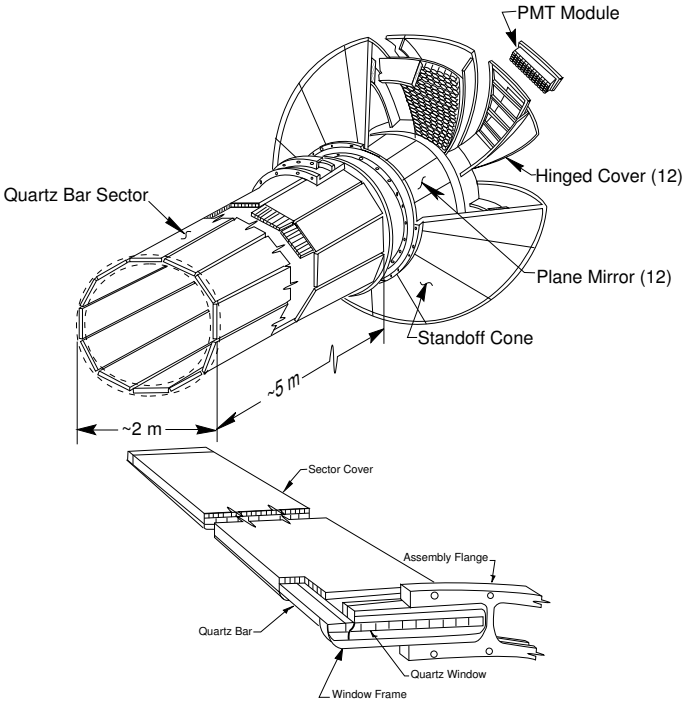

Figure 2.7. Mechanical elements of the *DIRC* and schematic view of bars assembled into a mechanical and optical sector.

Since particles are produced mainly forward in the detector because of the boost, the DIRC photon detector is placed at the backward end: the principal components of the DIRC are shown in Figure 2.7. The DIRC is placed in the barrel region and consists of 144 long, straight bars arranged in a 12-sided polygonal barrel. The bars are 1.7 cm-thick, 3.5 cm-wide and 4.90 m-long: they are placed into 12 hermetically sealed containers, called *bar boxes*, made of very thin aluminum-hexcel panels. Within

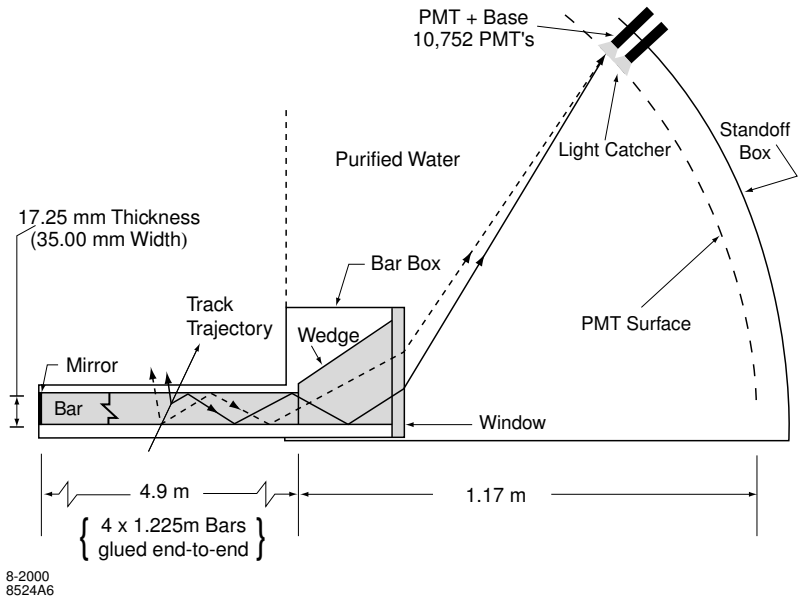

Figure 2.8. Schematics of the DIRC fused silica radiator bar and imaging region. Not shown is a 6 mrad angle on the bottom surface of the wedge.

a single bar box, 12 bars are optically isolated by a $\sim 150\,\mu m$ air gap enforced by custom shims made from aluminum foil.

The radiator material used for the bars is synthetic fused silica: the bars serve both as radiators and as light pipes for the portion of the light trapped in the radiator by total internal reflection. Synthetic silica has been chosen because of its resistance to ionizing radiation, its long attenuation length, its large index of refraction, its low chromatic dispersion within its wavelength acceptance.

The Cherenkov radiation is produced within these bars and is brought, through successive total internal reflections, in the backward direction outside the tracking and magnetic volumes: only the backward end of the bars is instrumented. A mirror placed at the other end on each bar reflects forward-going photons to the instrumented end. The Cherenkov angle at which a photon was produced is preserved in the propagation, modulo some discrete ambiguities (the forward-backward ambiguity can be resolved by the photon arrival-time measurement, for example). The DIRC

efficiency grows together with the particle incidence angle because more light is produced and a larger fraction of this light is totally reflected. To maximize the total reflection, the material must have a refractive index (fused silica index is $n = 1.473$) higher than the surrounding environment (the DIRC is surrounded by air with index $n = 1.0002$).

Once photons arrive at the instrumented end, most of them emerge into a water-filled expansion region (see Figure 2.8), called the *Standoff Box*: the purified water, whose refractive index matches reasonably well that of the bars ($n_{H_2O} = 1.346$), is used to minimize the total internal reflection at the bar-water interface.

The standoff box is made of stainless steel and consists of a cone, cylinder and 12 sectors of photomultipliers (PMT): it contains about 6000 liters of purified water. Each of the 12 PMT sectors contains 896 PMTs in a close-packed array inside the water volume: the PMTs are linearly focused – with a diameter of 2.9 cm – photomultiplier tubes, lying on an approximately toroidal surface.

The DIRC occupies only 8 cm of radial space, that allows for a relatively large radius for the drift chamber while keeping the volume of the CsI Calorimeter reasonably low: it corresponds to about 17% X_0 at normal incidence. The angular coverage is the 94% of the ϕ azimuthal angle and the 83% of $\cos\theta_{CM}$.

Cherenkov photons are detected in the visible and near-UV range by the PMT arrays. A small piece of fused silica with a trapezoidal profile glued at the back end of each bar allows for significant reduction in the area requiring instrumentation because it folds one half of the image onto the other half. The PMTs are operated directly in water and are equipped with light concentrators: the photomultiplier tubes are about $1.2\,m$ away from the end of the bars. This distance from the bar end to the PMTs, together with the size of the bars and PMTs, gives a geometric contribution to the single photon Cherenkov angle resolution of about 7 mrad. This is a bit larger than the resolution contribution from Cherenkov light production (mostly a 5.4 mrad

chromatic term) and transmission dispersions. The overall single photon resolution expected is about 9 mrad.

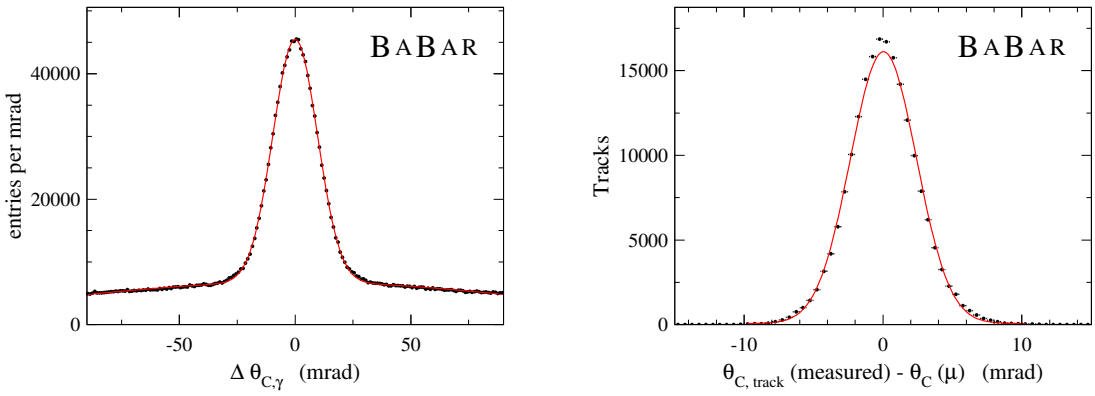

Figure 2.9. From di-muon data events, left plot: single photon Cherenkov angle resolution. The distribution is fitted with a double-Gaussian and the width of the narrow Gaussian is 9.6 mrad. Right plot: reconstructed Cherenkov angle for single muons. The difference between the measured and expected Cherenkov angle is plotted and the curve represents a Gaussian distribution fit to the data with a width of 2.4 mrad.

The image from the Cherenkov photons on the sensitive part of the detector is a cone cross-section whose opening angle is the Cherenkov angle modulo the refraction effects on the fused silica-water surface. In the most general case, the image consists of two cone cross-sections out of phase one from the other by a value related to an angle that is twice the particle incidence angle. In order to associate the photon signals with a track traversing a bar, the vector pointing from the center of the bar end to the center of each PMT is taken as a measure of the photon propagation angles α_x, α_y and α_z. Since the track position and angles are known from the tracking system, the three α angles can be used to determine the two Cherenkov angles θ_C and ϕ_C. In addition, the arrival time of the signal provides an independent measurement of the propagation of the photon and can be related to the propagation angles α. This over-

constraint on the angles and the signal timing are useful in dealing with ambiguities in the signal association and high background rates.

The expected number of photo-electrons (N_{pe}) is ~ 28 for a $\beta = 1$ particle entering normal to the surface at the center of a bar and increases by over a factor of of two in the forward and backward directions.

The time distribution of real Cherenkov photons from a single event is of the order of $50\,ns$ wide and during normal data taking they are accompanied by hundreds of random photons in a flat background distribution within the trigger acceptance window. The Cherenkov angle has to be determined in an ambiguity that can be up to 16-fold: the goal of the reconstruction program is to associate the correct track with the candidate PMT signal with the requirement that the transit time of the photon from its creation in the bar to its detection at the PMT be consistent with the measurement error of about $1.5\,ns$.

The resolution ($\sigma_{C,\mathrm{track}}$) on the track Cherenkov angle is expected to scale as

$$\sigma_{C,\mathrm{track}} = \frac{\sigma_{C,\gamma}}{\sqrt{N_{pe}}}$$

where $\sigma_{C,\gamma}$ is the single photon angle resolution. This angular resolution (obtained from di-muon events) can be estimated to be about 10.2 mrad, in good agreement with the expected value (see left plot in Figure 2.9). The measured time resolution is $1.7\,ns$ close to the intrinsic $1.5\,ns$ time spread of the PMTs.

2.5 Electromagnetic calorimeter

The electromagnetic calorimeter is designed to measure electromagnetic showers with excellent efficiency and energy and angular resolution over the energy range from 20 MeV to 9 GeV. This capability should allow the detection of photons from π^0 and η decays as well as from electromagnetic and radiative processes. The upper bound of the energy range is given by the need to measure QED processes like $e^+e^- \to e^+e^-(\gamma)$

and $e^+e^- \to \gamma\gamma$ for calibration and luminosity determination. The lower bound is set by the need for highly efficient reconstruction of B-meson decays containing multiple π^0s and η^0s. The measurement of very rare decays containing π^0s in the final state (for example, $B^0 \to \pi^0\pi^0$) puts the most stringent requirements on energy resolution, expected to be of the order of $1-2\%$. For energies below 2 GeV, the π^0 mass resolution is dominated by the energy resolution, while at higher energies, the angular resolution becomes dominant and it is required to be of the order of few mrad. The EMC is also used for electron identification and for completing the IFR output on μ and K_L^0 identification. It also has to operate in a $1.5\,T$ magnetic field.

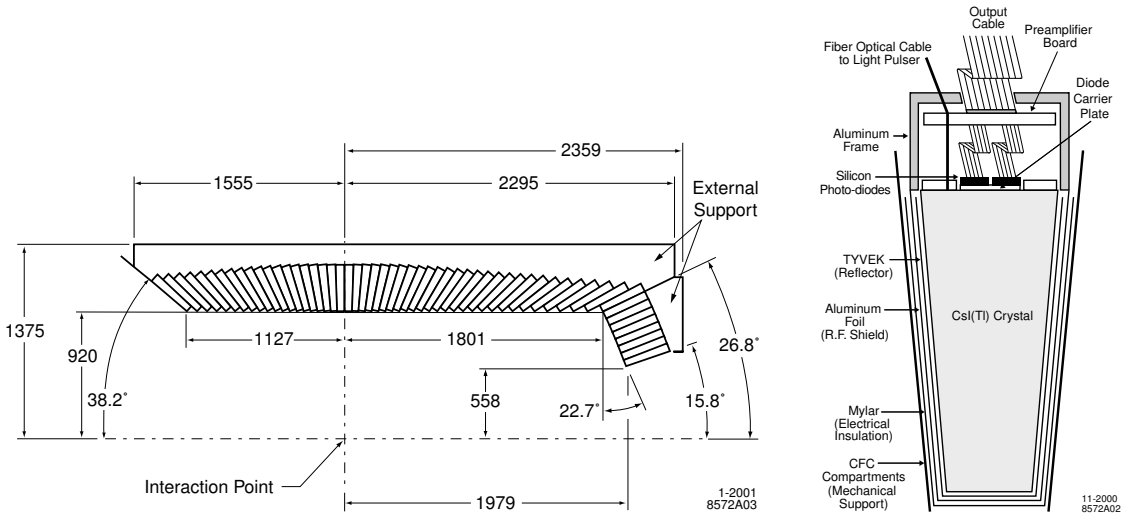

Figure 2.10. The electromagnetic calorimeter layout in a longitudinal cross section and a schematic view of the wrapped CsI(Tl) crystal with the front-end readout package mounted on the rear face (not to scale).

The EMC has been chosen to be composed of a finely segmented array of thallium-doped cesium iodide (CsI(Tl)) crystals. The crystals are read out with silicon photo-diodes which are matched to the spectrum of scintillation light. The energy resolution of a homogeneous crystal calorimeter can be described empirically in terms of a sum of two terms added in quadrature:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt[4]{E(\,\mathrm{GeV})}} \oplus b$$

where E and σ_E refer to the energy of a photon and its rms error, measured in GeV. The energy dependent term $a(\sim 2\%)$ arises basically from the fluctuations in photon statistics, but also from the electronic noise of the photon detector and electronics and from the beam-generated background which leads to large numbers of additional photons. This first term dominates at low energy, while the constant term $b(\sim 1.8\%)$ is dominant at higher energies (> 1 GeV). It derives from non-uniformity in light collection, leakage or absorption in the material in front of the crystals and uncertainties in the calibration.

The angular resolution is determined by the transverse crystal size and the distance from the interaction point: it can be empirically parameterized as a sum of an energy dependent and a constant term

$$\sigma_\theta = \sigma_\phi = \frac{c}{\sqrt{E(\,\mathrm{GeV})}} + d$$

where E is measured in GeV and with $c \sim 4$ mrad and $d \sim 0$ mrad.

In CsI(Tl), the intrinsic efficiency for the detection of photons is close to 100% down to a few MeV, but the minimum measurable energy in colliding beam data is about 20 MeV for the EMC: this limit is determined by beam and event-related background and the amount of material in front of the calorimeter. Because of the sensitivity of the π^0 efficiency to the minimum detectable photon energy, it is extremely important to keep the amount of material in front of the EMC at the lowest possible level.

Thallium-doped CsI has high light yield and small Molière radius in order to allow for excellent energy and angular resolution. It is also characterized by a short radiation length for shower containment at *BABAR* energies. The transverse size of the crystals is chosen to be comparable to the Molière radius achieving the required

angular resolution at low energies while limiting the total number of crystals and readout channels.

The *BABAR* EMC (left plot in Figure 2.10) consists of a cylindrical barrel and a conical forward end-cap: it has a full angle coverage in azimuth while in polar angle it extends from 15.8° to 141.8° corresponding to a solid angle coverage of 90% in the center of mass frame. Radially the barrel is located outside the particle ID system and within the magnet cryostat: the barrel has an inner radius of 92 cm and an outer radius of 137.5 cm and it is located asymmetrically about the interaction point, extending 112.7 cm in the backward direction and 180.1 cm in the forward direction. The barrel contains 5760 crystals arranged in 48 rings with 120 identical crystals each: the end-cap holds 820 crystals arranged in eight rings, adding up to a total of 6580 crystals. They are truncated-pyramid CsI(Tl) crystals (right plot in Figure 2.10): they are tapered along their length with trapezoidal cross-sections with typical transverse dimensions of 4.7×4.7 cm^2 at the front face, flaring out toward the back to about 6.1×6.0 cm^2. All crystals in the backward half of the barrel have a length of 29.6 cm: toward the forward end of the barrel, crystal lengths increase up to 32.4 cm in order to limit the effects of shower leakage from increasingly higher energy particles. All end-cap crystals are of 32.4 cm length. The barrel and end-cap have total crystal volumes of $5.2\,m^3$ and $0.7\,m^3$, respectively. The CsI(Tl) scintillation light spectrum has a peak emission at $560\,nm$: two independent photodiodes collect this scintillation light from each crystal. The readout package consists of two silicon PIN diodes, closely coupled to the crystal and to two low-noise, charge-sensitive preamplifiers, all enclosed in a metallic housing.

A typical electromagnetic shower spreads over many adjacent crystals, forming a *cluster* of energy deposit: pattern recognition algorithms have been developed to identify these clusters and to discriminate single clusters with one energy maximum from merged clusters with more than one local energy maximum, referred to as *bumps*.

The algorithms also determine whether a bump is generated by a charged or a neutral particle. Clusters are required to contain at least one seed crystal with an energy above 10 MeV: surrounding crystals are considered as part of the cluster if their energy exceeds a threshold of 1 MeV or if they are contiguous neighbours of a crystal with at least 3 MeV signal. The level of these thresholds depends on the current level of electronic noise and beam-generated background.

A bump is associated with a charged particle by projecting a track to the inner face of the calorimeter: the distance between the track impact point and the bump centroid is calculated and if it is consistent with the angle and momentum of the track, the bump is associated with this charged particle. Otherwise it is assumed to originate from a neutral particle.

On average, 15.8 clusters are detected per hadronic event: 10.2 are not associated to any charged particle. Currently, the beam-induced background contributes on average with 1.4 neutral clusters with energy above 20 MeV.

At low energy, the energy resolution of the EMC is measured directly with a 6.13 MeV radioactive photon source (a neutron-activated fluorocarbon fluid) yielding $\sigma_E/E = 5.0 \pm 0.8\%$. At high energy, the resolution is derived from Bhabha scattering where the energy of the detected shower can be predicted from the polar angle of the electrons and positrons. The measured resolution is $\sigma_E/E = 1.9 \pm 0.1\%$ at 7.5 GeV. Figure 2.11 shows the energy resolution on data compared with expectations from Monte Carlo. From a fit to the experimental results to Eq. (2.5), $a = 2.32 \pm 0.30\%$ and $b = 1.85 \pm 0.12\%$ are obtained. The constant term comes out to be greater than expected: this is mainly caused by a cross talk effect, still not corrected, in the front-end electronics.

The measurement of the angular resolution is based on Bhabha events and ranges between 12 mrad and 3 mrad going from low to high energies. A fit to Eq. (2.5) results in $c = (3.87 \pm 0.07)$ mrad and $d = (0.00 \pm 0.04)$ mrad.

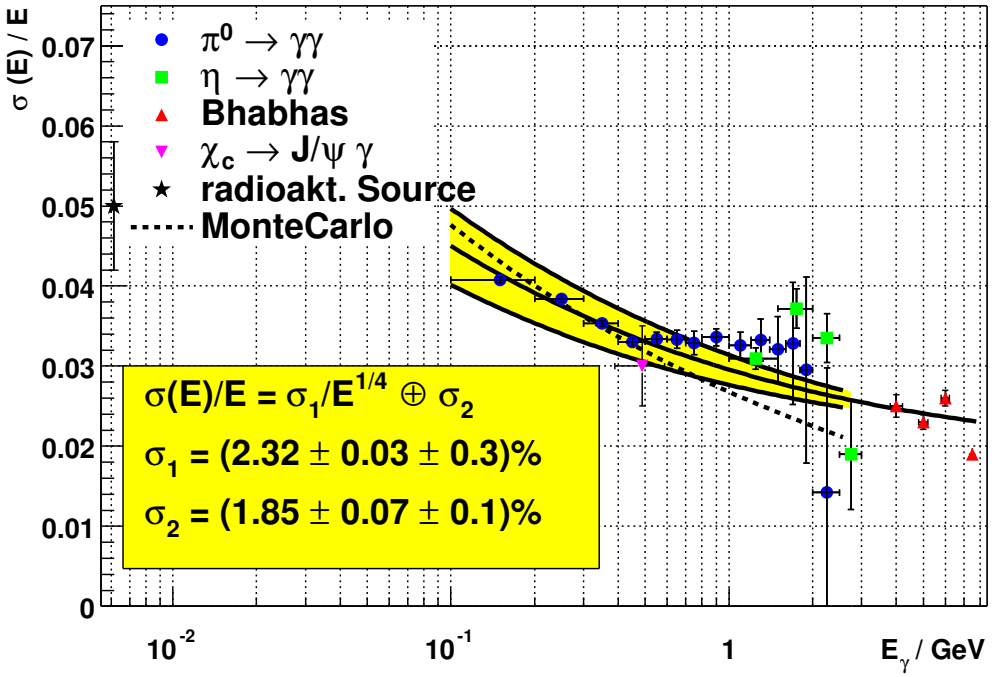

Figure 2.11. EMC resolution as a function of the energy.

2.6 Instrumented flux return

The Instrumented Flux Return detector is dedicated to muon identification and neutral hadrons detection (mainly K_L^0) in a wide range of momentum and angles.

The IFR, as all the other *BABAR* subsystems, has an asymmetric structure with a polar angle coverage that is $17° \leq \theta_{lab} \leq 150°$. The IFR (Figure 2.12) was originally made of 19 layers of Resistive Plate Chambers (RPC) in the barrel region and 18 layers in forward and backward regions, placed inside the iron layers used for the solenoidal magnetic field return yoke. Recently, part of the RPCs have been replaced by Limited Streamer Tubes (LST).

The iron structure is subdivided in three main parts: the barrel one surrounding the solenoid, made of 6 sextants covering the radial distance between 1.820 m and 3.045 m with a length of 3.750 m (along the z axis); the forward end-cap and backward end-cap covering the forward (positive z axis) and backward regions. The endcaps

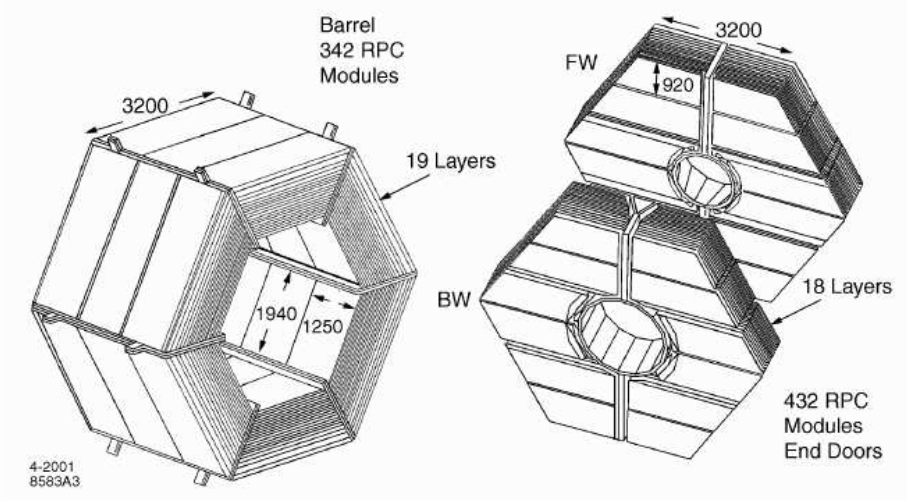

Figure 2.12. IFR view

are instrumented with RPCs, the barrel is completely instrumented with LSTs since 2006.

The RPC section is shown in Figure 2.13.

Signals produced by particles crossing the gas gap inside the RPCs are collected on both sides of the chamber by using thin strips (thickness $\sim 40\,\mu m$) with witdh of the order of a centimeter. Strips are applied in two orthogonal directions on insulating planes 200 μm thick, in order to have a bi-dimensional view.

The used gas mixture is made of 56.7% Argon, 38.8% Freon-134a and 4.5% Isobutane. Working voltage for RPCs is $\sim 7.5\ kV$. Iron layers keeping apart RPC planes are chilled by a water system that keeps the temperature $\sim 20^oC$. RPC efficiencies have been measured by using cosmics taken on a weekly base.

The gaps in the IFR barrel are filled with LSTs. Each tube is composed by 7 or 8 cells. Each cell has a $18 \times 18\,mm$ section with a silver plated wire at the center and a resistive graphite coating. The wire provides the high voltage while the signal is read by strips like in RPCs. The operating voltage is typically 5500 V. A gas admixture of 2.5% Argon, 9.5% Isobutane and 88% CO_2 is used.

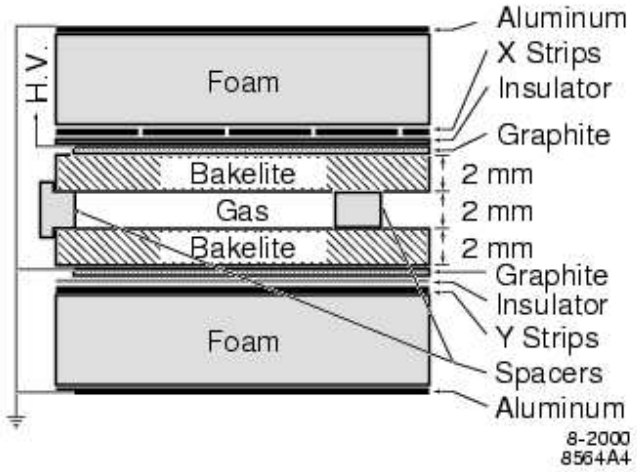

Figure 2.13. RPC section with HV connection scheme.

Muons are identified by measuring the number of traversed interaction lengths in the entire detector and comparing it with the number of expected interaction lengths for a muon of a given momentum. Moreover, the projected intersections of a track with the RPC planes are computed and, for each readout plane, all strips clusters detected within a predefined distance from the predicted intersection are associated with the track: the average number and the r.m.s. of the distribution of RPC and LST strips per layer gives additional μ/π discriminating power. We expect in fact the average number of strips per layer to be larger for pions producing a hadronic interaction than for muons. Other variables exploiting clusters distribution shapes are constructed. Selection criteria based on all these variables are applied to select muons.

CHAPTER 3

SEARCH FOR $B^0 \to \ell^+\ell'^-$ DECAYS

3.1 Analysis overview

The event selection for this analysis is very similar to another *BABAR* analysis: $B^0 \to h^+h'^-$ [56], where the B meson decays into pions and kaons. We use the same reconstruction code and initial event selection criteria, except for the leptonic vetoes (see Section 4.3).

The selected sample is characterized by the presence of three main sources of background: QED events (leptons produced in e^+e^- collisions), continuum $q\overline{q}$ events and $B \to hh$ events with charged pions and kaons misidentified as leptons. In order to suppress the QED background, we reject events with less than five tracks, whereas for the suppression of the other backgrounds we use the event topology and the standard muon and electron selectors: neural network (NN) based selectors for muons and likelihood based selectors for electrons (the choice of the selectors is discussed in sec. 3.4). We use these particle identity (PID) selectors to distinguish $B \to ll$ from $B \to hh$ events in different samples. We perform four independent fits to these components: the fit to the hh sample – to estimate the shape of $q\bar{q}$ background probability density functions (PDF's) – and the three fits to the leptonic samples – to estimate the corresponding yields.

Selection efficiency associated to the PID requirements is evaluated on Monte Carlo events, after applying data/Monte Carlo corrections obtained from dedicated samples and collected in tables.

In the fit to the hh sample, we use the information on the Cherenkov angle – obtained from the DIRC – to distinguish two types of $B\bar{B}$ events: $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$.[1] For consistency, DIRC requirements are also applied to the leptonic samples, in order not to bias the shapes of the $q\bar{q}$ background (because the QED background, which affects the leptonic samples but not the hh one, is more abundant in the forward and backward region, where the DIRC information is not available).

In addition, because of the parametric dependence of the DIRC PDF on the track momentum, we distinguish three $q\bar{q}$ background components ($b\pi\pi$, $bK\pi$, and bKK), such that the normalization of $B \to hh$ and $q\bar{q}$ components are the same. The events associated to the $bK\pi$ (bKK) components are those having one (two) tracks identified as kaons by the DIRC PDF. All the other events (including $q\bar{q}$ events with real leptons) are associated to the $b\pi\pi$ component. We found that only this component has a significative cross-feed in the leptonic samples (see appendix A), so that the DIRC PDF is not used in the fits to extract the three leptonic yields. In addition, $B \to hh$ events do not represent a significant source of background after PID requirements are applied. The contribution of other $B\bar{B}$ background is found to be negligible, as described in section 3.5.1.

3.2 Datasets

This analysis is based on data taken with the *BABAR* detector during Run 1 through 5, corresponding to $\sim 3.8 \times 10^8$ B-meson pairs (equivalent to an integrated luminosity of 347 fb^{-1} at the $\Upsilon(4S)$ peak). We use Monte Carlo samples to optimize selection criteria. Each signal MC sample consists of 290×10^3 events. The following exclusive modes have been considered:

[1] Even if we consider $B^0 \to K^+K^-$ events in the analysis optimization, we found the contamination from these events negligible (this mode is still unobserved) and we ignore it when performing the fit.

- $B^0 \to \mu^+\mu^-$,
- $B^0 \to e^\pm\mu^\mp$,
- $B^0 \to e^+e^-$,
- $B^0 \to \pi^+\pi^-$,
- $B^0 \to K^+K^-$,
- $B^0 \to K^+\pi^-$.

To ensure that any source of continuum background is taken into account in the background characterization, we avoid to use Monte Carlo samples for the continuum. Instead, we use the sidebands of the kinematic variables $m_{\rm ES}$ and ΔE (both variables described in Section 3.6.1.1) on actual data:

- **on-resonance data:** 383.6 ± 4.2 million of $B\bar{B}$ events (corresponding to $\sim$ 347 fb^{-1});
- **off-resonance data:** 36.8 fb^{-1} of data taken 40 MeV below the $\Upsilon(4S)$ resonance;

corresponding to the entire Run1 - Run5 dataset. The study of possible $B\bar{B}$ background contamination is performed on the following decays:

- $B \to \pi l\nu$: 309×10^3 events,
- $B \to K^{*+}\pi^-$: 145×10^3 events,
- $B \to K^{*0}\pi^+$: 154×10^3 events,
- $B \to K^+\pi^-\pi^+$: 618×10^4 events,
- $B \to \rho^+\pi^-$: 1455×10^3 events,
- $B \to \rho^0\pi^+$: 158×10^3 events.

3.3 Event selection

Reconstructed candidates are composed of two charged tracks with the criteria listed in Table 3.1, assuming pion mass hypothesis. To further reduce the contami-

Select tracks with	Selection criteria
distance in $x-y$ plane	$\|d_{xy}\| < 1.5$ cm
distance in z axis	$\|d_z\| < 10$ cm
minimum number of DCH hits	$N_{\rm DCH} > 12$
minimum track fit χ^2	$\chi^2 > 0$
maximum momentum	$p_{\rm lab} < 10$ GeV
minimum transverse momentum	$p_\perp > 0.1$ GeV
geometrical acceptance	$0.410 < \theta_{\rm lab} < 2.54$ rad

Table 3.1. Summary of track selection criteria.

nation of continuum background, we apply the following requirements:

- $\mathbf{R}_2 < 0.95$

  The Fox-Wolfram moments [63] H_n are defined as:

$$H_n = \sum_{i,j} \frac{|\mathbf{p}_i|.|\mathbf{p}_j|}{E_{\rm vis}^2} P_n(cos\theta_{ij}), \tag{3.1}$$

  where P_n are the Legendre polynomials, $\mathbf{p}_{i,j}$ are the momenta of particles i and j, θ_{ij} is the angle between the direction of these particles, and $E_{\rm vis}$ is the total visible energy of the event.

  If we neglect the particle masses, energy and momentum conservation requires $H_0 = 1$, while for two-jet events:

  - $H_1 = 0$;
  - $H_l \simeq 1$ for even values of l;
  - $H_l \simeq 0$ for odd values of l.

R_2 is the ratio of the second to zeroth Fox-Wolfram moment: $R_2 = \frac{H_2}{H_0}$. It is largest for continuum events where the particle momenta are aligned along one axis. The distribution of R_2 computed with all the particles in the event.

- **Sphericity** > 0.01

  We define a second order tensor, called *sphericity tensor*:

$$T_{\alpha\beta} = \sum_j (\delta_{\alpha\beta} p_j^2 - p_{j\alpha} p_{j\beta}) \tag{3.2}$$

  Since it is symmetric, it has three real eigenvalues. The sphericity is the minimum of the three eigenvalues and can be written as:

$$S = \min_{\hat{n}} S(\hat{n}) = \min_{\hat{n}} \frac{3\sum_{i=1}^{N} p_{i\perp}^{*2}}{2\sum_{i=1}^{N} p_i^{*2}} \tag{3.3}$$

  The sum is over all charged particles in the event. $\hat{n}$ is a generic unit vector and $p_{i\perp}^*$ is the component of the momentum of particle i perpendicular to $\hat{n}$, calculated in the $\Upsilon(4S)$ frame. The sphericity axis is the direction $\hat{n}$ which satisfies equation 3.3, *i.e.* the direction that minimizes the sum of the momentum components perpendical to it. The unit vector $\hat{n}$ is the eigenvector of $T_{\alpha\beta}$ associated to S.

- $|\cos\theta_S| < 0.8$

  θ_S is the angle between the sphericity axes of the B candidate and of the rest of the event.

- **Level 3 tag bit**

  We require information from at least one of the two Level 3 trigger outputs from the DCH and of the EMC.

- Events belonging to the fit region: $|\Delta E| < 0.150$ GeV and $5.2 < m_{\rm ES} < 5.2895$ GeV/c^2.

  The *energy difference* ΔE is defined as

$$\Delta E = E_B^* - \sqrt{s}, \tag{3.4}$$

  where E_B^* is the energy of the B candidate in the $\Upsilon(4S)$ rest frame – i.e. the center of mass frame (CM) – and $\sqrt{s}$ is the total energy of the e^+e^- system in the CM frame. The resolution on this variable is affected by the detector momentum resolution and by the particle identification since a wrong mass assignment implies a shift in ΔE.

  The *beam-energy substituted mass* $m_{\rm ES}$ is defined as:

$$m_{\rm ES} = \sqrt{(\sqrt{s}/2)^2 - p_B^{*2}}, \tag{3.5}$$

  where p_B^* is the B candidate momentum in the CM frame. Since $|p_B^*| \ll \sqrt{s}/2$, the experimental resolution on $m_{\rm ES}$ is dominated by beam energy fluctuations. These two kinematic variables are in practice uncorrelated, because they have different sources of experimental smearing (detector momentum resolution for ΔE and energy of the beams for $m_{\rm ES}$).

- **Fisher** discriminant (defined in section 3.6) in the range $[-3, 3]$

In order to have meaningful information from the DIRC, we require the Cherenkov angle to be $\theta_c > 0$ and we apply a cut on the minimum number of photons, $N_\gamma > 5$. Hence, our event sample is free from pure background hits on the DIRC which might be reconstructed as tracks, so that the θ_c pull can be calculated. [2] We remove eventual

[2]Evaluating the pull, the expected value of θ_c is calculated from the momentum of the track, for a given mass hypothesis (pion or kaon), the corresponding PDF for the leptonic channels is the same as $\pi\pi$.

protons by applying an outlier cut on θ_c, by requiring that the corresponding pull is not rejected by a 4σ cut in at least one of the mass hypotheses.

All the above cuts have been optimized for $B \to hh$ analysis. We do not repeat the optimization study, since the features of signal and background in our dataset are the same. Moreover, the largest part of the selection involves the rest of the event, which is independent on the channel studied for the reconstructed B.

The requirement of a minimum of four tracks is used to remove QED background, as it is explained in Section 3.5.3.

3.3.1 Bremsstrahlung recovery

The six plots in Figure 3.1 show the distribution of ΔE for the six Monte Carlo samples.

The distribution of the reconstructed energy of the B candidate in signal Monte Carlo shows a significative asymmetry, due to the fact that the energy of the photons produced by QED final state radiation is not associated to the tracks in the usual track reconstruction. This effect is expected to be relevant more for electrons than for muons. In fact, we see a larger tail contribution for $e\mu$ and ee Monte Carlo samples.

In order to improve the energy resolution, we try to recover the energy lost with the standard *BABAR* algorithm for *bremsstrahlung recovery*. For each track, we look if there is a photon associated to it; if that is the case, we add the photon to the event. In this way, we reduce the size of the radiative tail in ΔE, obtaining a better energy resolution and increasing the rejection power against continuum background. This is illustrated in Figure 3.1, where we show the comparison of ΔE distribution for signal Monte Carlo samples with and without the bremsstrahlung recovery algorithm: in the case of final states with one or more electrons, we observe a significant migration of events from the tail to the peak of the distribution, while no significant improvement is obtained in the case of $\mu\mu$ and hh samples.

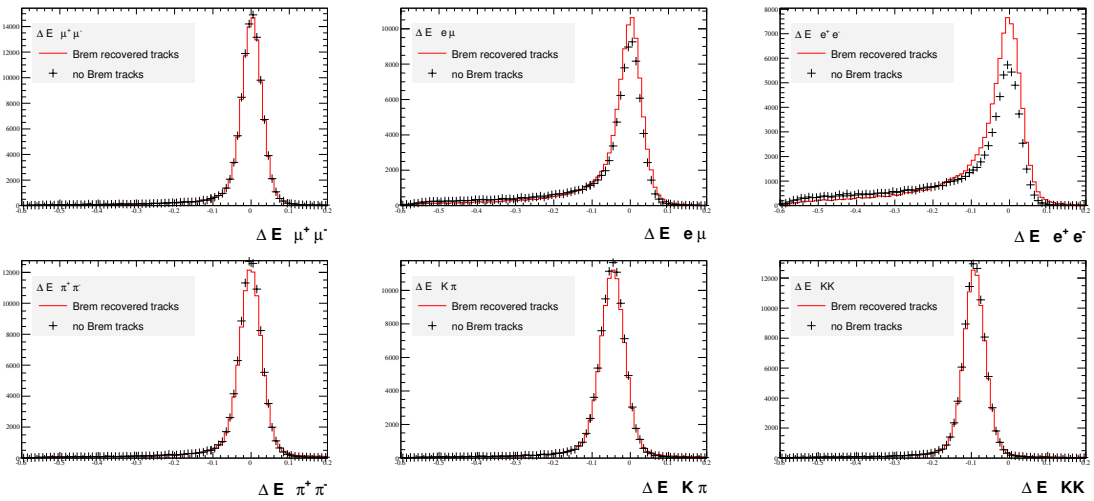

Figure 3.1. ΔE distribution for signal Monte Carlo samples of $\mu^+\mu^-$ (top-left), $e^\pm\mu^\mp$ (top-center), e^+e^- (top-right), $\pi^+\pi^-$ (bottom-left), $K^\pm\pi^\mp$ (bottom-center) and K^+K^- (bottom-right), with (histogram) or without (points) using the *bremsstrahlung recovery*. All the tracks are reconstructed assuming pion mass hypothesis.

3.3.2 Signal and fit regions

In the previous *BABAR* analysis [57] the effect of energy loss on the ΔE shape has been taken into account by defining asymmetric fit regions in this variable and optimizing different regions for different decay channels. On the other hand, our fit strategy is based on a Maximum Likelihood (ML) fit. We decided to adopt a common ΔE window, choosing it in such a way that all the different signal species can be determined and that in any case we have enough sideband events to determine background shapes. In practice, we use the fit region of $B \to hh$ analysis, which is

- $5.2\ \mathrm{GeV}/c^2 < m_{ES} < 5.2895\ \mathrm{GeV}/c^2$,
- $|\Delta E| < 150$ MeV.

When optimizing the PID requirements, we reduce this region to a signal box that takes into account the core of signal events and only the fraction of $q\bar{q}$ background component that falls under the signal peaks. We use the largest box of the previous *BABAR* $B \to ll$ analysis:

	$\mu\mu$	$e\mu$	ee	$\pi\pi$	$K\pi$
Reconstruction	82.1 ± 0.1	78.9 ± 0.1	74.5 ± 0.1	78.8 ± 0.1	76.6 ± 0.1
number of tracks > 4	80.9 ± 0.1	80.9 ± 0.1	81.1 ± 0.1	81.3 ± 0.1	80.9 ± 0.1
`BGFMultiHadron`	$100. \pm 0.$	$100. \pm 0.$	$100. \pm 0.$	$100. \pm 0.$	$100. \pm 0.$
`Level 3`	$100. \pm 0.$	$100. \pm 0.$	99.9 ± 0.0	$100. \pm 0.$	$100. \pm 0.$
R_2	$100. \pm 0.$	$100. \pm 0.$	$100. \pm 0.$	$100. \pm 0.$	$100. \pm 0.$
sphericity	99.9 ± 0.0	99.9 ± 0.0	99.9 ± 0.0	99.9 ± 0.0	99.9 ± 0.0
`GoodTracksAccLoose`	87.4 ± 0.1	74.5 ± 0.1	63.3 ± 0.1	87.1 ± 0.1	87.5 ± 0.1
$\Delta E - m_{ES}$	95.8 ± 0.0	88.8 ± 0.1	81.1 ± 0.1	94.9 ± 0.1	94.7 ± 0.1
$\lvert\cos\theta_S\rvert$	78.5 ± 0.1	78.4 ± 0.1	78.3 ± 0.1	75.8 ± 0.1	76.4 ± 0.1
`Fisher`	$100. \pm 0.$	$100. \pm 0.$	$100. \pm 0.$	$100. \pm 0.$	$100. \pm 0.$
Total $B \to hh$ selection	43.6 ± 0.1	33.0 ± 0.1	24.2 ± 0.1	40.1 ± 0.1	39.2 ± 0.1
θ_c	81.7 ± 0.1	82.3 ± 0.1	82.5 ± 0.1	81.8 ± 0.1	81.6 ± 0.1
N_γ	97.4 ± 0.0	97.5 ± 0.1	97.8 ± 0.1	96.1 ± 0.1	95.5 ± 0.1
θ_c outlier	99.7 ± 0.0	99.3 ± 0.0	$99. \pm 0.$	97.7 ± 0.0	97.8 ± 0.0
Total PID selection	79.3 ± 0.0	79.7 ± 0.0	$79.9 \pm 0.$	76.8 ± 0.0	76.1 ± 0.0
Total efficiency	34.6 ± 0.1	26.3 ± 0.1	19.4 ± 0.1	30.8 ± 0.1	29.8 ± 0.1

Table 3.2. Summary of selection efficiencies (%) for several exclusive modes, estimated using Monte Carlo samples (290k generated events for each mode). The BGFMultiHadron tag-bit consists of two charged tracks satisfying a very loose requirement on R_2. GoodTracksAccLoose summarizes the selection criteria described in Table 3.1.

- $5.274\ \mathrm{GeV}/c^2 < m_{\mathrm{ES}} < 5.2895\ \mathrm{GeV}/c^2$,
- $-105\ \mathrm{MeV} < \Delta E < 55\ \mathrm{MeV}$,

corresponding to the e^+e^- channel in the study of Ref. [57].

3.3.3 Efficiencies

Using $B\bar{B}$ Monte Carlo samples, we calculate the selection efficiency. The efficiencies of the individual cuts, applied in sequence, are summarized in Table 3.2. For the sake of completeness, we also report in this table the efficiency of the cut on the number of tracks, which is described in Section. 3.5.3.

3.4 PID optimization

We choose the best PID selection among several combinations of muon and electron selectors. This is a crucial point of the analysis, since the PID selectors are the only handle we have to isolate samples containing the leptonic yields (2μ, $1\mu 1e$, and $2e$ in the following) from a sample dominated by $q\bar{q}$ background (used for background studies).

In each case, we compute the number of expected events for continuum background and for each of the five B decay modes ($\pi\pi$, $K\pi$, $\mu\mu$, $e\mu$, and ee). Since we apply both electron and muon selectors on both the tracks, we have 16 different possibilities (four possible results for each selector). We look at the statistical significance $\sigma = N_S/\sqrt{N_S + N_B}$ (where N_S is the number of signal events for a given decay and N_B is the number of background $q\bar{q}$ events) and we try to merge these 16 categories in a smaller number of categories, such that each one is dominated by a single signal component. We also optimize the selectors in order to maximize the statistical significance of the dominant component in each category.

In order to evaluate N_S, we estimate signal efficiencies from Monte Carlo, assuming the following decay rates:

- $BR(B^0 \to \mu^+\mu^-) = 8.3 \cdot 10^{-6}$,
- $BR(B^0 \to e^\pm\mu^\mp) = 1.8 \cdot 10^{-7}$,
- $BR(B^0 \to e^+e^-) = 6.1 \cdot 10^{-8}$,
- $BR(B^0 \to \pi^+\pi^-) = 4.8 \cdot 10^{-6}$,
- $BR(B^0 \to K^+\pi^-) = 1.8 \cdot 10^{-5}$.

The first three values correspond to the 90% C.L. upper limits quoted by the most recent BABAR analysis [57] and the last two are taken from the PDG [50]. All the samples are scaled to $205\,\mathrm{fb}^{-1}$, corresponding to the Run 1-4 luminosity.

In order to estimate σ in the signal region, we need to know the number of signal and background events in the signal box. For signal, this is directly calculated from the ratio of selection efficiencies in the signal-box and in the entire fit region.

For continuum background, we extrapolate it from the number of events in the sidebands (Run1-Run4 dataset), scaling it according to the relative areas of the background distributions: we fit the shape of $m_{\rm ES}$ (ΔE) for events taken from the sideband of ΔE ($m_{\rm ES}$) and we use the output distributions to obtain N_B. The sidebands are defined as:

- $5.2 < m_{\rm ES} < 5.24$ GeV/c^2,
- $150 < |\Delta E| < 300$ MeV.

The fits, shown in the plots of Figure 3.2, give the following scaling factors (ratio of the area of the signal region over the area of the sideband):

- $R_{\Delta E} = 0.53$
- $R_{m_{\rm ES}} = 0.39$

The number of continuum events N_B in the signal region is therefore:

$$N_B = N_{sb} \cdot R_{\Delta E} \cdot R_{m_{\rm ES}} \tag{3.6}$$

where N_{sb} is the number of events in the sideband of the on-resonance data.

We define four PID requirements, which identify four non-overlapping samples (hh, 2μ, $1e1\mu$, and $2e$), that associate each possible combination of PID requirements to one (and only one) of them. When the two tracks pass only one selector, it is easy to assign the event to a category. For instance, an event with both the tracks passing only electron selector will be identified as a $2e$ event without any ambiguity.

On the other hand, additional care is needed for intermediate cases:

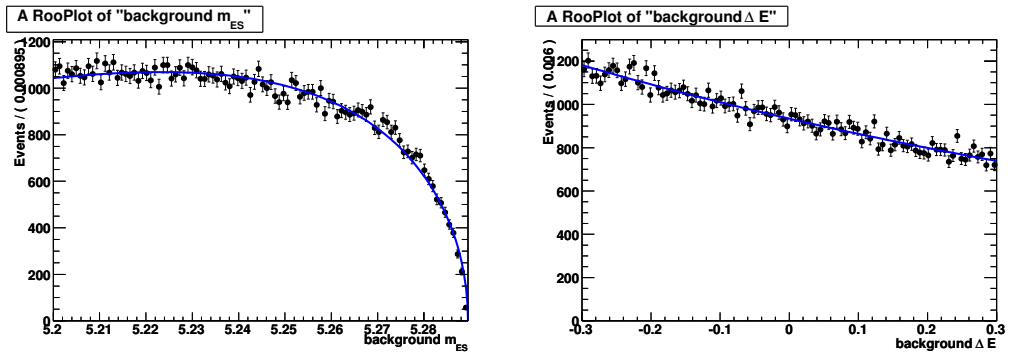

Figure 3.2. Fit of $m_{\rm ES}$ (left) ΔE (right) shapes of on-resonance data (Run1-Run4 dataset), assuming an ARGUS [60] function (an exponential) for $m_{\rm ES}$ (ΔE). The $m_{\rm ES}$ (ΔE) distribution is for events taken from the sideband of ΔE (m_{ES})

- one track passes a selector and the other one fails both;
- an event with at least one track that passes both electron and muon selectors.

In order to associate these events to one of the four categories we need to look for the number of expected signal and $q\bar{q}$ background events for each possibility and find the case with the optimal statistical significance and signal efficiency.

3.4.1 Optimal selectors

We restrict our search of the optimal selectors to the following possibilities:

- muons: Neural Network based selectors, with loose, tight, and very tight requirements;
- electrons: likelihood based.

For each selector we count the number of events for all the decays (3 leptonic and 2 hadronic) in the signal region and the number of $q\bar{q}$ background events as described in section 3.4. All the yields are rescaled to $205\,\mathrm{fb}^{-1}$. The statistical significance for each decay is then calculated and reported in Table 3.3, Table 3.4, and Table 3.5. The rows in the tables represent the sixteen possible combinations of PID selectors,

whereas we report in the columns the statistical significance for each reconstructed decay. The two hadronic decays are grouped into the more general decay $B^0 \to h^+h^-$, since they represent the same source of background in this analysis. For each of the three muon selectors, we report the number of on-resonance events into the entire fit region, as shown in Table 3.6.

NN loose muons, likelihood based electrons					
	$\mu^+\mu^-$	$e^\pm\mu^\mp$	e^+e^-	h^+h^-	$q\bar{q}$
Category	σ	σ	σ	σ	N_B
hh	0.0001	0.	0.	3.9	5707
$\mu\mu$	0.204	0.	0.	0.0009	13.1
$e^+\mu^-$	0.	0.15	0.	0.	4.
μ^+e^-	0.	0.13	0.	0.	4.4
ee	0.	0.	0.10	0.	6.8
μ^+h^-	0.0023	0.0001	0.	0.067	273
$h^+\mu^-$	0.0030	0.0001	0.	0.076	294
e^+h^-	0.	0.0022	0.0002	0.	60.1
h^+e^-	0.	0.0023	0.0002	0.	66.9
$(e,\mu)^+h^-$	0.	0.	0.	0.	0.
$h^+(e,\mu)^-$	0.	0.	0.	0.	0.
$(e,\mu)^+\mu^-$	0.	0.	0.	0.	0.
$\mu^+(e,\mu)^-$	0.	0.	0.	0.	0.
$e^+(e,\mu)^-$	0.	0.	0.	0.	0.
$(e,\mu)^+e^-$	0.	0.	0.	0.	0.
$(e,\mu)^+(e,\mu)^-$	0.	0.	0.	0.	0.

Table 3.3. Statistical significance ($\sigma = N_S/\sqrt{N_S + N_B}$) and number of $q\bar{q}$ events N_B table for the selectors NN loose for muons and likelihood based for electrons. The rows represent the sixteen possible categories corresponding to the chosen PID selectors, whereas the columns are the different samples. Statistical significance is calculated from the number of signal events (both hadronic and leptonic) taken from Monte Carlo samples and number of background events from the sideband of on-resonance data. We label as (e,μ) those tracks passing both electron and muon selectors.

We choose the following selectors as optimal:

- NN tight for muons;
- likelihood based for electrons.

NN tight muons, likelihood based electrons					
	$\mu^+\mu^-$	$e^\pm\mu^\mp$	e^+e^-	h^+h^-	$q\bar{q}$
Category	σ	σ	σ	σ	N_B
hh	0.0002	0.	0.	4.4	6022
$\mu\mu$	0.205	0.	0.	0.	4.6
$e^+\mu^-$	0.	0.16	0.	0.	2.3
μ^+e^-	0.	0.12	0.	0.	3.
ee	0.	0.	0.10	0.	6.8
μ^+h^-	0.0056	0.0001	0.	0.012	124
$h^+\mu^-$	0.0077	0.0001	0.	0.011	135
e^+h^-	0.	0.0047	0.0002	0.	61.9
h^+e^-	0.	0.0049	0.0002	0.	68.2
$(e,\mu)^+h^-$	0.	0.	0.	0.	0.
$h^+(e,\mu)^-$	0.	0.	0.	0.	0.
$(e,\mu)^+\mu^-$	0.	0.	0.	0.	0.
$\mu^+(e,\mu)^-$	0.	0.	0.	0.	0.
$e^+(e,\mu)^-$	0.	0.	0.	0.	0.
$(e,\mu)^+e^-$	0.	0.	0.	0.	0.
$(e,\mu)^+(e,\mu)^-$	0.	0.	0.	0.	0.

Table 3.4. Statistical significance ($\sigma = N_S/\sqrt{N_S + N_B}$) and number of $q\bar{q}$ events N_B table for the selectors NN tight for muons and likelihood based for electrons. The rows represent the sixteen possible categories corresponding to the chosen PID selectors, whereas the columns are the different samples. Statistical significance is calculated from the number of signal events (both hadronic and leptonic) taken from Monte Carlo samples and number of background events from the sideband of on-resonance data. We label as (e,μ) those tracks passing both electron and muon selectors.

NN very tight muons, likelihood based electrons					
	$\mu^+\mu^-$	$e^\pm\mu^\mp$	e^+e^-	h^+h^-	$q\bar{q}$
Category	σ	σ	σ	σ	N_B
h^+h^-	0.0006	0.	0.	4.5	6093
$\mu^+\mu^-$	0.17	0.	0.	0.	2.1
$e^+\mu^-$	0.	0.17	0.	0.	1.2
μ^+e^-	0.	0.095	0.	0.	2.7
e^+e^-	0.	0.	0.10	0.	6.8
μ^+h^-	0.0086	0.0001	0.	0.0065	90.9
$h^+\mu^-$	0.012	0.0001	0.	0.0070	99.8
e^+h^-	0.	0.0079	0.0002	0.	62.7
h^+e^-	0.	0.0081	0.0002	0.	68.6
$(e,\mu)^+h^-$	0.	0.	0.	0.	0.
$h^+(e,\mu)^-$	0.	0.	0.	0.	0.
$(e,\mu)^+\mu^-$	0.	0.	0.	0.	0.
$\mu^+(e,\mu)^-$	0.	0.	0.	0.	0.
$e^+(e,\mu)^-$	0.	0.	0.	0.	0.
$(e,\mu)^+e^-$	0.	0.	0.	0.	0.
$(e,\mu)^+(e,\mu)^-$	0	0.	0.	0.	0.

Table 3.5. Statistical significance ($\sigma = N_S/\sqrt{N_S + N_B}$) and number of $q\bar{q}$ events N_B table for the selectors NN very tight for muons and likelihood based for electrons. The rows represent the sixteen possible categories corresponding to the chosen PID selectors, whereas the columns are the different samples. Statistical significance is calculated from the number of signal events (both hadronic and leptonic) taken from Monte Carlo samples and number of background events from the sideband of on-resonance data. We label as (e,μ) those tracks passing both electron and muon selectors.

Category	loose	tight	very tight
h^+h^-	54765	57923	58572
$\mu^+\mu^-$	147	46	35
$e^+\mu^-$	37	24	20
μ^+e^-	53	36	28
e^+e^-	51	51	51
μ^+h^-	3155	1281	933
$h^+\mu^-$	2349	1166	876
e^+h^-	558	571	575
h^+e^-	594	611	619
$(e,\mu)^+h^-$	0	0	0
$h^+(e,\mu)^-$	0	0	0
$(e,\mu)^+\mu^-$	0	0	0
$\mu^+(e,\mu)^-$	0	0	0
$e^+(e,\mu)^-$	0	0	0
$(e,\mu)^+e^-$	0	0	0
$(e,\mu)^+(e,\mu)^-$	0	0	0

Table 3.6. Number of on-resonance data events in the fit region for each of the three NN criteria (loose, tight and very tight) for muon selection and likelihood selector for electrons, for runs 1 to 4, corresponding to 205 fb^{-1}. We label as (e,μ) those tracks passing both electron and muon selectors.

A description of these selection criteria is outlined in appendix A. Comparing the two tables we notice that tightening the muon selector there is a small migration of events from the leptonic categories to the categories with at least one hadron; therefore the statistical significance of the leptonic categories associated to the tight muon selector is slightly higher with respect to the very tight one.

3.4.2 Definition of the samples using PID

Imposing the PID requirements, we can distinguish the sixteen possibilities listed in the previous section. We group them into four independent samples:

- hh sample, composed of all the events having at least one track that fails both electron and muon selectors;
- 2μ sample, composed of all the events having both tracks which pass only the muon selector;
- $1e1\mu$ sample, composed of all the events having one track that passes only the muon selector and the other that passes only the electron selector;
- $2e$ sample, composed of all the events having both tracks which pass only the electron selector.

The few events which have one lepton and one hadron are included in the hh sample. In addition, we remove the very few tracks which pass both electron and muon selectors.

3.4.3 PID efficiency correction

To estimate the leptonic BFs, we need to know the efficiency of the PID requirements used to identify the four independent samples. The values of these efficiencies, obtained from signal Monte Carlo, are given in Table 3.7. We note that the electron identification is in general more efficient than the muon identification.

Sample	hh	2μ	$1e1\mu$	$2e$
$\mu^+\mu^-$	0.4995 ± 0.0016	0.5004 ± 0.0016	$(10.0 \pm 3.2) \cdot 10^{-5}$	0.
$e^\pm\mu^\mp$	0.3418 ± 0.0017	0.	0.6581 ± 0.0017	$(5.2 \pm 2.6) \cdot 10^{-5}$
e^+e^-	0.1445 ± 0.0015	0.	0.	0.8555 ± 0.0015
$K^+\pi^-$	0.99984 ± 0.00004	$(1.62 \pm 0.43) \cdot 10^{-4}$	0.	0.
$\pi^+\pi^-$	0.99924 ± 0.00009	$(7.17 \pm 0.90) \cdot 10^{-4}$	$(4.5 \pm 2.4) \cdot 10^{-5}$	0.

Table 3.7. Category efficiency for each sample, obtained counting the number of events of each Monte Carlo decay sample which fall in the given category.

The Monte Carlo simulation does not perfectly reproduce the measured performance of the PID selector. In order to take into account this data/MC discrepancy, we correct our Monte Carlo with standard *BABAR* tables obtained from ad hoc data samples and calculate the new efficiencies for each decay mode and each sample. The obtained parameters, listed in Table 3.8, will be used in the fit on data.

Comparing this result with Table 3.7, we notice that the efficiency of the hh category increases after applying the data/MC correction (i.e. the pure categories for the extraction of leptonic modes are less efficient than what the Monte Carlo predicts). This effect is more evident for the identification of muons, where the data control sample is composed of muons coming from $e^+e^- \rightarrow \mu^+\mu^-\gamma$. The biggest discrepancy between data and MC is seen in lower momentum muons.

The evaluation of the systematic error associated to the values of the efficiencies is described in Section 3.7.3.

Sample	hh	2μ	$1e1\mu$	$2e$
$\mu\mu$	0.5472 ± 0.0016	0.4527 ± 0.0016	$(11.0 \pm 3.3) \cdot 10^{-5}$	0.
$e\mu$	0.3491 ± 0.0017	0.	0.6508 ± 0.0017	$(5.2 \pm 2.6) \cdot 10^{-5}$
ee	0.1465 ± 0.0015	0.	$(5.3 \pm 3.1) \cdot 10^{-5}$	0.8535 ± 0.0015
$K\pi$	0.99988 ± 0.00004	$(1.1 \pm 0.4) \cdot 10^{-4}$	0.	0.
$\pi\pi$	0.99950 ± 0.00008	$(4.3 \pm 0.7) \cdot 10^{-4}$	$(7.8 \pm 3.0) \cdot 10^{-5}$	0.

Table 3.8. Category efficiency for each sample, obtained counting the number of events of each Monte Carlo decay sample which fall in the given category, using Monte Carlo corrected with the PID data/MC tables.

3.5 Background Characterization

This analysis is affected by three sources of background:

- $B\bar{B}$ background, originating from other B decays having two high momentum tracks. As discussed below, the only contribution of this type comes from $B \to h\pi$ ($h = \pi, K$) events, which are considered as a signal in the analysis and are fitted together with the leptonic signal yields.

- $q\bar{q}$ events ($q = u, d, s, c$), which is the dominant source of background for charmless analyses.

- QED events with pairs of high-momentum muons or electrons, which is a special source of background associated with the event topology we study.

Because of the different nature of these background sources, different strategies are used in order to control their effect on signal.

$B \to h\pi$ ($h = \pi, K$) events are very similar to our leptonic signal events, because of the presence of two hard tracks in the final state. The use of PID allows to reduce this contamination (see Section 3.4), requiring that the two tracks agree with the considered final state hypothesis ($\mu\mu$, $e\mu$, ee). On the other hand, the PID requirements do not have in general a 100% efficiency in removing this background. In addition, the $q\bar{q}$ background (see Section 3.5.2) is so high that a residual contamination survives the PID requirements. The residual events will not be enough to model the shape of the background PDF (which will introduce systematic errors from the knowledge of the shapes) and too many to be ignored. A better choice is to isolate hadronic events from leptonic events and study the shapes of the $q\overline{q}$ component there.

This picture is valid only if QED background can be removed, since it introduces differences in shape depending on the PID requirements.

3.5.1 $B\bar{B}$ background

The $B\bar{B}$ background can affect our fit in two ways: the estimate of $q\bar{q}$ PDF's shape in the fit of the hh sample, and the determination of the signal yield in the leptonic samples.

In the hh fit, the presence of $B\bar{B}$ events wrongly identified as $q\bar{q}$ events might bias the result of the fit. The main contribution comes from $B \to hh$ events, which are explicitly taken into account in the fit by adding specific likelihood components for $B \to \pi\pi$ and $B \to K\pi$. We neglect the additional contribution from the (still unobserved) $B \to KK$ processes. Other possible sources of $B\bar{B}$ background in this sample have been studied in detail in the time-dependent analysis of $B^0 \to h^+h'^-$ [56]. The only relevant contribution, namely $B \to \rho\pi$ events with one pion lost in the reconstruction, is neglected in the time-dependent analysis, because of the very low reconstruction efficiency of the ΔE cut. In fact, being one particle in the final state lost in the reconstruction, ΔE is shifted to negative values. Rather than ignoring this component, we remove it by imposing $\Delta E > -100$ MeV. This is not an acceptable solution for the $B \to hh$ analysis, since it significatively affect the efficiency of $B \to K\pi$ events. On the other hand, it is perfect in our case, since we are only interested to the shape of background events. In principle a problem is induced by the fact that we have to extrapolate the shape of ΔE from the range [-100 MeV, 150 MeV] back to the range [-150 MeV, 150 MeV]. This is not a serious issue, since the shape of ΔE is almost linear. As a matter of fact, using the extra requirement $\Delta E > -100$ MeV when fitting for $q\bar{q}$ background shapes and using specific components of the likelihood for $B \to hh$ events, the presence of $B\bar{B}$ background in the hh samples does not affect our result.

On the other hand, the $B\bar{B}$ background might have an impact on the determination of the leptonic yields. In order to identify possible sources of background, we apply our selection algorithm on a sample of inclusive $B\bar{B}$ decays. On a statistics

		Fit region				Signal Box			
Decay Mode	$BR[10^{-6}]$	hh	2μ	$1\mu 1e$	$2e$	hh	2μ	$1\mu 1e$	$2e$
$B \to \pi l \nu$	134	81.296	1.583	1.007	0.000	7.914	0.288	0.000	0.000
$B \to K^{*+}\pi^-$	9.8	163.389	0.000	0.000	0.000	0.359	0.000	0.000	0.000
$B \to K^{*0}\pi^+$	10.7	0.025	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$B \to K^+\pi^-\pi^+$	1.8(UL)	0.048	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$B \to \rho^+\pi^-$	24.0	694.031	0.050	0.047	0.000	6.399	0.000	0.000	0.000
$B \to \rho^0\pi^+$	8.7	154.181	0.009	0.000	0.000	0.822	0.000	0.000	0.000

Table 3.9. Number of $B\bar{B}$ background events in the four PID categories, evaluated on exclusive Monte Carlo samples. They are rescaled to an integrated luminosity of 347 fb^{-1}.

corresponding to 1050 fb^{-1}, we only find one event in the $1\mu 1e$ sample, while no candidates survive our selection for 2μ and $2e$ samples. Using the information of the Monte Carlo truth available in our ntuples, we can only say that the two particles are two pions. The values of the fit variables ($m_{ES} = 5.2761$ GeV/c^2, $\Delta E = -0.083$ GeV, and Fisher = -1.24611) suggest that this event would be identified as a $q\bar{q}$ background candidate, being on the tails of ΔE and Fisher and $\sim 1\sigma$ far from the m_{ES} peak.

A more detailed study was done using the list of *dangerous* channels identified for $B \to hh$ analysis and applying our selection on the exclusive Monte Carlo samples. In Table 3.9 we list the expected number of events for these channels in our four samples. The tables display the yield in the fit region, as well as in the signal box, for an integrated luminosity corresponding to 347 fb^{-1}.

All these studies suggest that any source of additional $B\bar{B}$ background (besides $B \to hh$ events) does not affect our analysis.

3.5.2 $q\bar{q}$ background

The separation between $q\bar{q}$ background events and $B\bar{B}$ decays is achieved exploiting the different decay topology: on one hand, $q\bar{q}$ events exhibit a jet-like structure, the two jets oriented back-to-back along the line of flight of the two quarks origi-

nally produced in the e^+e^- collision. On the other hand, final state particles in $B\bar{B}$ decays are isotropically distributed in the $\Upsilon(4S)$ rest frame. In order to reduce the contamination, we require $R_2 < 0.95$, sphericity > 0.01, and $|\cos\theta_S| < 0.8$, as already described in sec. 4.3. In addition, we define the two quantities L_2 and L_0 as

$$L_j \equiv \sum_i |\mathbf{p}_i^*||\cos\theta_i^*|^j, \tag{3.7}$$

and we build a Fisher [59] discriminant from them. The coefficients of the linear combination have been optimized on a large sample of $B\bar{B}$ and $q\overline{q}$ simulated events. The use of this variable and the kinematic variables $m_{\rm ES}$ and ΔE in the fit (see Section 3.6) allows to discriminate the residual $q\overline{q}$ component from $B\bar{B}$ events.

3.5.3 QED background

The QED background comes from pairs of charged leptons (muons or electrons) produced by the colliding beams. These events are characterized by a low number of tracks. This can be seen comparing the distribution of the number of tracks in Figure 3.3 for $q\overline{q}$ Monte Carlo, on-resonance sidebands and off-resonance in the leptonic PID categories (contaminated by QED background) and hh (where no QED background is present).

The distributions of the number of tracks for the three signal Monte Carlo events are shown in Figure 3.4. They follow the same shape of events belonging to hh category. Table 3.10 shows the mean number of these tracks and the standard error on the mean.

These QED events are characterized by a distribution in $m_{\rm ES}$ which deviates from a phase space shape. This is shown in Figure 3.5, where the distribution of $m_{\rm ES}$ for events with less (left) or more (right) than four tracks are shown. The distributions of ΔE and of the Fisher discriminant do not show the effect ot the QED events, as can be seen in Figure 3.6.

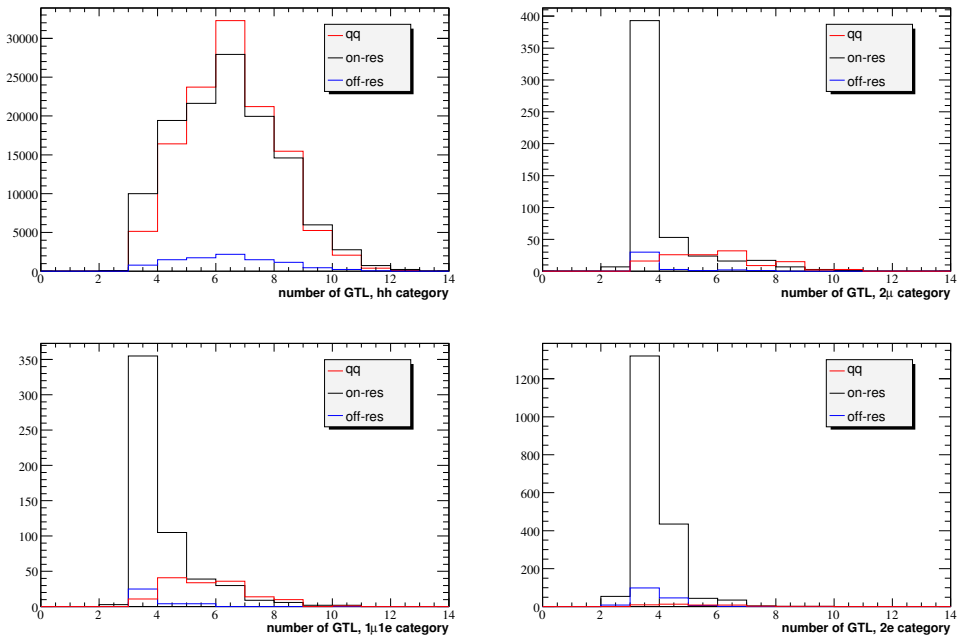

Figure 3.3. Distribution of the number of tracks per event in PID category hh (top-left) 2μ (top-right), $1\mu 1e$ (bottom-left) and $2e$ (bottom-right), for $q\bar{q}$ (red), on-resonance ΔE sidebands (black) and off-resonance (blue).

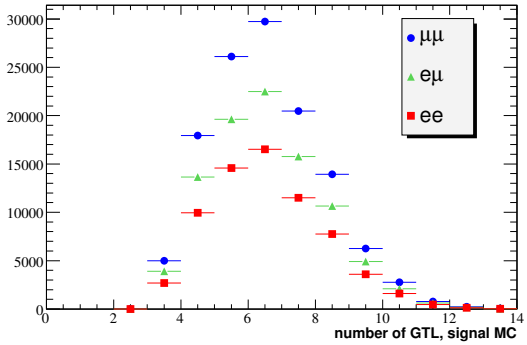

Figure 3.4. Distribution of the number of tracks per events for the three leptonic Monte Carlo samples: $B^0 \to \mu\mu$ (circle), $B^0 \to e\mu$ (triangle) and $B^0 \to ee$ (square).

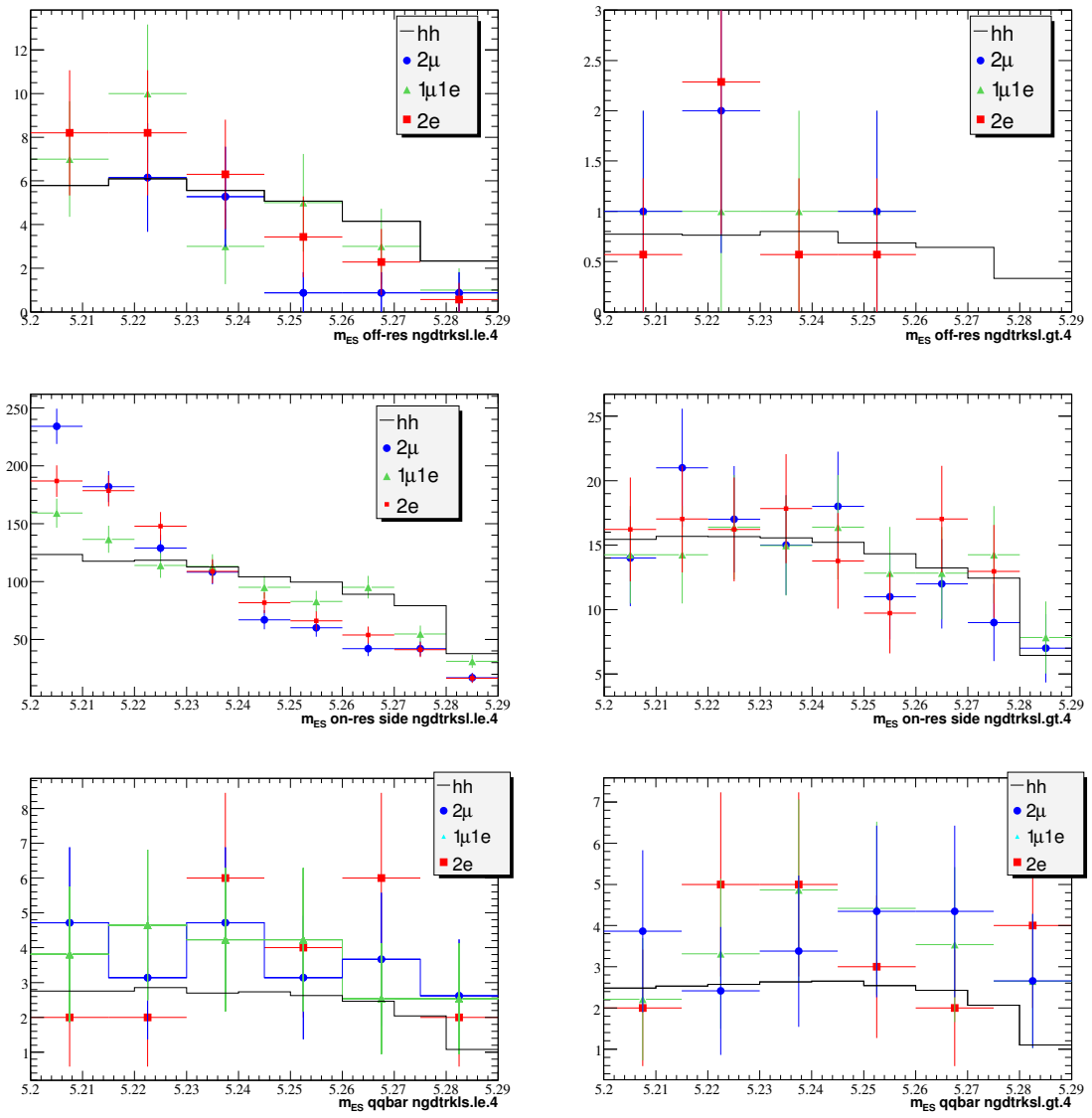

Figure 3.5. Distribution of $m_{\rm ES}$ for events with less than five (left) or more than four (right) tracks, from samples of off-resonance (top), on-resonance ΔE sidebands (middle), and $q\overline{q}$ Monte Carlo (bottom).

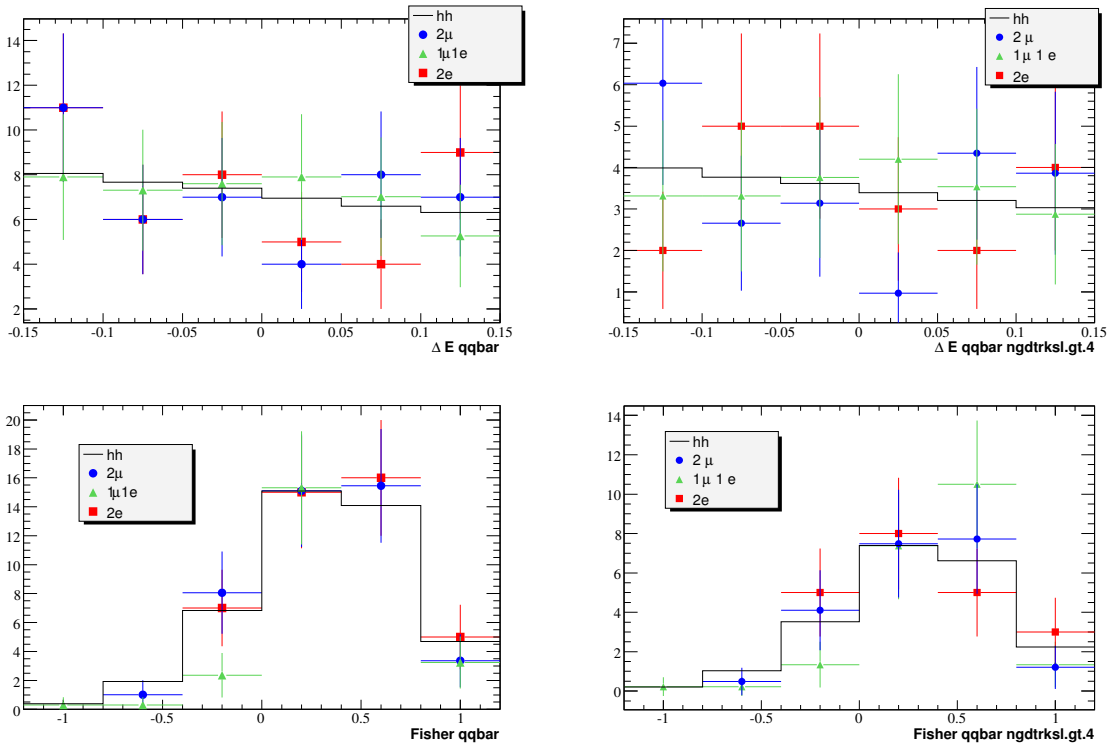

Figure 3.6. Distribution of ΔE (top) and of the Fisher discriminant (bottom) for all events (left) and for events with more than four (right) tracks, from $q\bar{q}$ Monte Carlo (bottom).

Dataset	Mean number of tracks
$\mu\mu$	6.055 ± 0.005
$e\mu$	6.063 ± 0.006
ee	6.062 ± 0.006
$\pi\pi$	6.576 ± 0.005
$K\pi$	6.059 ± 0.005
on-res sideband hh	5.943 ± 0.005
on-res sideband 2μ	3.518 ± 0.053
on-res sideband $1\mu 1e$	3.657 ± 0.050
on-res sideband $2e$	3.320 ± 0.016
off-res hh	5.937 ± 0.018
off-res 2μ	3.40 ± 0.16
off-res $1\mu 1e$	3.36 ± 0.12
off-res $2e$	3.333 ± 0.052

Table 3.10. Mean number of tracks and standard error on the mean for the five Monte Carlo samples ($\mu\mu$, $e\mu$, ee, $\pi\pi$ and $K\pi$), for the sideband of the on-resonance and for the off-resonance Run1 - Run5 data, with events which belong to the PID categories hh, 2μ, $1\mu 1e$ and $2e$.

While for hh category we do not observe significative differences among data and continuum Monte Carlo, a deviation from ARGUS [60] shape is observed in the leptonic categories for those events with less than five tracks. In order to remove this difference (which makes it impossible to extrapolate from hh category the cross-feed of non $B\bar{B}$ background in leptonic categories), we require the number of tracks in the event to be greater than four (which has an efficiency of $\sim 80\%$ on signal samples, reducing the QED background to negligible levels).

3.6 Maximum likelihood fit

We extract yields of the three signal leptonic species through a chain of fits, performed on different data samples.

We divide the sample into four samples, using the output of muon and electron PID selectors (as explained in Section 3.4). The first step is a fit of the hh sample, done to determine the shapes of the background PDF's, to be used in the nomi-

nal fits of 2μ, $1e1\mu$, and $2e$ samples. Since hh includes $B \to \pi\pi$, $B \to \pi K$, and $B \to KK$ events, and almost all the continuum background events ($\sim 99\%$ of the total $q\overline{q}$ background yield), while only few leptonic events are expected, we can safely neglect the presence of leptonic signal events and fit for hadronic yields and $q\overline{q}$ background. This assumption induces a systematic on the background estimation. For instance, neglecting the presence of $B \to \mu\mu$ events in the hh category implies that $N^{2\mu}_{\mu\mu}(1-\epsilon^{2\mu}_{\mu\mu})/\epsilon^{2\mu}_{\mu\mu}$ events will be assigned to a wrong component of the likelihood when performing the fit to hh. Nevertheless, considering the fact that the hadronic and continuum components have typically hundreds of events, we will need to associate a systematic error to this hypothesis only in presence of a significative leptonic signal (*i.e.* a significative enhancement with respect to the SM).

Once the previous fit is done, we perform a fit to $N_{\mu\mu}$ in the 2μ subsample, in which all the parameters (except the signal and continuum background yields) are fixed to what comes out of the previous fit. The same procedure is independently applied to $1e1\mu$ and $2e$ subsamples.

As a validation, we then determine the relative fractions of continuum background events in 2μ, $1e1\mu$, and $2e$ samples, performing a simultaneous fit of the four PID on-resonance data sidebands (see appendix A). The sidebands are defined requiring $m_{\rm ES} < 5.25$ GeV/c^2 and $\Delta E > 50$ MeV, which rejects signal and $B\bar{B}$ background contributions. Knowing the fractions and the yield of continuum background in the hh sample (first fit), we can calculate the amount of continuum background events and compare that to what we determine from the fit.

All the fits are executed maximizing an unbinned maximum likelihood (ML). The likelihood functions are defined using $m_{\rm ES}$, ΔE and the $Fisher$ discriminant. The fit on the the hh sample also uses the Cherenkov angle θ_c. In the following, we describe in the details all the fits.

3.6.1 Fit of the hh subsample

We perform an unbinned multidimensional ML fit of the hh sample to extract the yields of $\pi\pi$, $K\pi$, and $q\overline{q}$ events. We use the Cherenkov angles θ_c^+ and θ_c^- of the positive and negative charged tracks to separate kaons from pions. For background, Cherenkov angle is also added to the likelihood, in order to have the same overall normalization for the signal and background components of the likelihood. This technical detail implies that we fit for three different background yields ($b\pi\pi$, $bK\pi$, and bKK) and two signal yields ($\pi\pi$, and $K\pi$). [3] In addition, distinguishing for the charge of the kaon, we are simultaneously sensitive to the two *CP* conjugated states $K^+\pi^-$ and $K^-\pi^+$. Rather than fitting for the two yields, we parameterize them as a function of the *CP*-averaged yield and the *CP* asymmetry (both for signal and background). Because of this, we have to introduce two additional parameters: $A_{K\pi}$, and $A_{bK\pi}$. The likelihood function is based on $m_{\rm ES}$, ΔE, $Fisher$, and θ_c. The full expression is:

$$
\begin{aligned}
\mathcal{L} \quad = \quad & \frac{e^{-(N_{\pi\pi}+N_{K\pi}+N_{b\pi\pi}+N_{bK\pi}+N_{bKK})/N}}{\sqrt[N]{(N_{\pi\pi}+N_{K\pi}+N_{b\pi\pi}+N_{bK\pi}+N_{bKK})!}} \cdot \prod_{i=1}^{N} \Big\{ \\
N_{\pi\pi}\cdot \quad & P_{\pi\pi}(m_{ES})_i \cdot P_{\pi\pi}(\Delta E)_i \cdot P_\pi^+(\theta_c^+)_i \cdot P_\pi^-(\theta_c^-)_i \cdot P_{hh}(\mathcal{F})_i + \\
N_{K\pi}\cdot \quad & \frac{1-A_{K\pi}}{2} \cdot P_{\pi\pi}(m_{ES})_i \cdot P_{K\pi}(\Delta E)_i \cdot P_K^+(\theta_c^+)_i \cdot P_\pi^-(\theta_c^-)_i \cdot P_{hh}(\mathcal{F})_i + \\
N_{K\pi}\cdot \quad & \frac{1+A_{K\pi}}{2} \cdot P_{\pi\pi}(m_{ES})_i \cdot P_{K\pi}(\Delta E)_i \cdot P_\pi^+(\theta_c^+)_i \cdot P_K^-(\theta_c^-)_i \cdot P_{hh}(\mathcal{F})_i + \\
N_{b\pi\pi}\cdot \quad & P_{bhh}(m_{ES})_i \cdot P_{bhh}(\Delta E)_i \cdot P_\pi^+(\theta_c^+)_i \cdot P_\pi^-(\theta_c^-)_i \cdot P_{bhh}(\mathcal{F})_i + \\
N_{bK\pi}\cdot \quad & \frac{1-A_{bK\pi}}{2} \cdot P_{bhh}(m_{ES})_i \cdot P_{bhh}(\Delta E)_i \cdot P_K^+(\theta_c^+)_i \cdot P_\pi^-(\theta_c^-)_i \cdot P_{bhh}(\mathcal{F})_i + \\
N_{bK\pi}\cdot \quad & \frac{1+A_{bK\pi}}{2} \cdot P_{bhh}(m_{ES})_i \cdot P_{bhh}(\Delta E)_i \cdot P_\pi^+(\theta_c^+)_i \cdot P_K^-(\theta_c^-)_i \cdot P_{bhh}(\mathcal{F})_i + \\
N_{bKK}\cdot \quad & P_{bhh}(m_{ES})_i \cdot P_{bhh}(\Delta E)_i \cdot P_K^+(\theta_c^+)_i \cdot P_K^-(\theta_c^-)_i \cdot P_{hh}(\mathcal{F})_i \Big\}, \qquad (3.8)
\end{aligned}
$$

[3] Since the dedicate $B^0 \to h^+h'^-$ analysis did not find any evidence of $B \to KK$ decays, we neglect this component in our fit.

where:

- N is the total number of events observed in the fitted sample;
- $N_{\pi\pi}$ and $N_{K\pi}$ are the yields of the hadronic $B\bar{B}$ background;
- $N_{bK\pi}$, N_{bKK} and $N_{b\pi\pi}$ are the yields of the $q\overline{q}$ background with or without one of the two tracks identified as a kaon;
- $P(X)_i$ is the PDF of the observable X for the event i;
- $A_{K\pi}$ and $A_{bK\pi}$ are the CP asymmetries for signal and $q\overline{q}$ background.

The $P(X)_i$ term labels the PDF of the observable X for the event i. We distinguish P^+ and P^- for the DIRC, since the corresponding PDF parameterization – obtained from the $D^* \to D\pi$ control sample – is charge sensitive. We give below the detailed description of how the P functions are modeled in the fit.

This fit provides the values of the parameters of continuum background PDF's. These values are used in the fit to $B \to ll$ samples (see Section 3.6.2), while the error on the parameters is used to estimate the systematic error on the leptonic yields. It also provides the mean values of $m_{\rm ES}$ and ΔE distributions for $B \to \pi\pi$ events. The difference among these values and those obtained from a fit to $B \to \pi\pi$ Monte Carlo samples is used to correct the mean values of $m_{\rm ES}$ and ΔE distributions of $B \to ll$ components of the nominal fit (see Section 3.6.2).

In addition, by floating the signal shapes one by one, we determine the scaling factors for $B \to ll$ shapes, which are used to evaluate the systematic error on the leptonic yields (see Section 3.7.1).

In order to suppress any contribution from $B\bar{B}$ background, we require $\Delta E > -100$ MeV. We verified that the values obtained for the background shapes are consistent to those obtained with a looser cut $\Delta E > -150$ MeV (as expected, since the $B\bar{B}$ background component is known to be small).

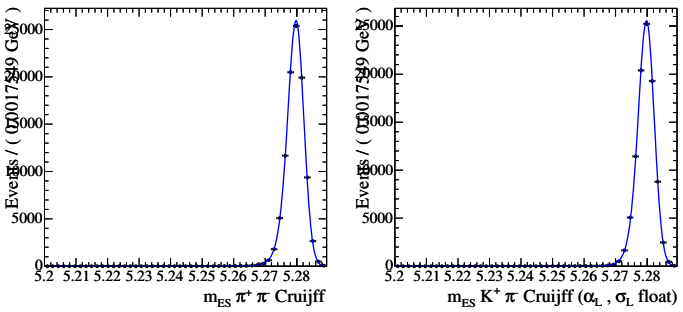

Figure 3.7. Fit to $m_{\rm ES}$ distributions for $\pi^+\pi^-$ (left) and $K^+\pi^-$ (right), obtained from signal Monte Carlo samples.

3.6.1.1 Parameterization of $B \to hh$ PDF's

We parameterize $B \to hh$ PDF's using unbinned ML fits to a sample of signal Monte Carlo events, after applying the PID weighting to improve the data/Monte Carlo agreement.

$B\bar{B}$ events peak near the B mass in the $m_{\rm ES}$ distributions and are described using a Cruijff function

$$f(x) = e^{-\frac{(x-\mu)^2}{2\cdot\sigma^2_{R/L}+\alpha_{R/L}\cdot(x-\mu)^2}}, \tag{3.9}$$

where μ is the maximum of the distribution, σ is the RMS of the Gaussian component and α is a parameter describing non-Gaussian tails. The parameters α and σ are different in the case $x > \mu$ (α_R and σ_R) and $x < \mu$ (α_L and σ_L). Since the right side of the distribution is a Gaussian, we fix $\alpha_R = 0$ when determining the parameters of the function from a set of Monte Carlo samples. Figure 3.7 shows the distribution of the $m_{\rm ES}$ shapes for $B \to \pi\pi$ and $B \to K\pi$. We do not observe any deviation between the two shapes. In the ML fit, we will assume the same shapes for the two PDF's, obtained from a fit to the $\pi\pi$ sample; in addition, the mean of the Cruijff function is floated in this fit, whereas all the other parameters are fixed.

The shape of ΔE is described with a Cruijff function. Because of the wrong mass hypothesis, $K\pi$ distribution is negatively shifted. The distribution of $B \to hh$ Monte

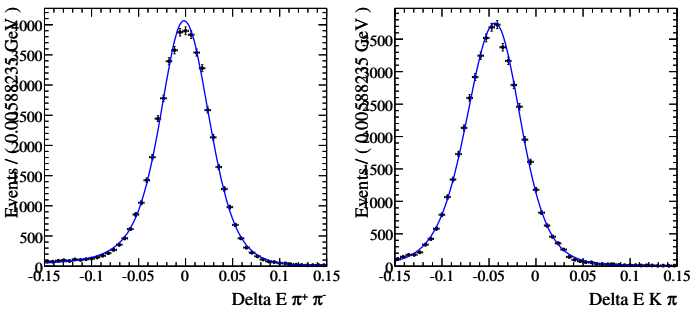

Figure 3.8. Fit to ΔE distributions for $\pi^+\pi^-$ (left) and $K^+\pi^-$ (right), obtained from unbinned fits to signal Monte Carlo samples.

Carlo events are shown in Figure 3.8. In the fit to data, the mean value is floated, while all the other parameters are fixed to the Monte Carlo value.

Discrimination of signal from background in the ML fit is enhanced by the use of a Fisher discriminant $\mathcal{F}$. We use the two-variable Legendre Fisher used in the $B^0 \to h^+h'^-$ analysis. It is constructed from the CM momentum p_i and angle θ_i of each particle i in the rest of the event (r.o.e.) with respect to the thrust axis of the B candidate. Specifically, we define $\mathcal{F}$ as:

$$\mathcal{F} = 0.5319 - 0.6023L_0 + 1.2698L_2, \tag{3.10}$$

where $L_0 = \sum_i^{\text{r.o.e.}} p_i$ and $L_2 = \sum_i^{\text{r.o.e.}} p_i \left|\cos\theta_i\right|^2$. The coefficients come from a training of variable on a sample of pure signal vs pure background events. We use a bifurcated Gaussian to describe the $B\bar{B}$ distribution. The shapes are shown in Figure 3.9.

The parameters used for these PDF's in the nominal fit to the hh data sample are summarized in Table 3.11.

Separation of pions from kaons is achieved using the pion and kaon θ_C pulls for each of the tracks. The pulls are described using double Gaussian functions, as in $B^0 \to h^+h^-$ analysis [56]. The parameters are known functions of track momenta (as obtained from a $D^* \to D\pi$ control samples) and are different for positive ($P_\pi^+(\theta_c^+)$

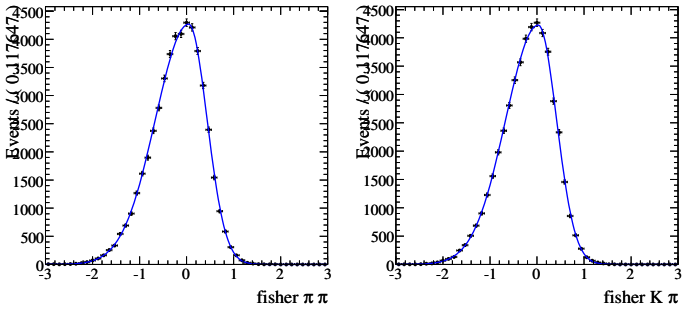

Figure 3.9. Fit to $\mathcal{F}$ distributions for $\pi^+\pi^-$ (left) and $K^+\pi^-$ (right), obtained from an unbinned fit to signal Monte Carlo samples.

Variable	$K^+\pi^-$	$\pi^+\pi^-$
SIG $m_{\rm ES}$ α_L	0.1014 ± 0.0011	0.1014 ± 0.0011
SIG $m_{\rm ES}$ σ_L	0.002532 ± 0.000016	0.00253 ± 0.00002
SIG $m_{\rm ES}$ α_R	0.	0.
SIG $m_{\rm ES}$ σ_R	0.002481 ± 0.000014	0.00248 ± 0.00001
SIG ΔE α_L	0.1095 ± 0.0049	0.074 ± 0.043
SIG ΔE σ_L	0.02857 ± 0.00038	0.0298 ± 0.0014
SIG ΔE α_R	0.1003 ± 0.0022	0.0068 ± 0.0055
SIG ΔE σ_R	0.02571 ± 0.00027	0.02531 ± 0.00095
SIG $\mathcal{F}$ μ	-0.0044 ± 0.0044	-0.0044 ± 0.0044
SIG $\mathcal{F}$ σ_L	0.6832 ± 0.0030	0.6832 ± 0.0030
SIG $\mathcal{F}$ σ_R	0.3835 ± 0.0025	0.3835 ± 0.0025

Table 3.11. Summary of the fit parameters, obtained with 290k events of Monte Carlo of the two hadronic decays considered in the fits.

and $P_K^+(\theta_c^+)$ of Eq. 4.3) and negative ($P_\pi^-(\theta_c^-)$ and $P_K^-(\theta_c^-)$ of Eq. 4.3) charged tracks. They are common to $B\bar{B}$ and $q\overline{q}$ candidates.

3.6.1.2 $q\overline{q}$ background PDF's

The distribution of $m_{\rm ES}$ background events follow the phase space curve described by the ARGUS function [60]

$$f(x) = N \cdot x \cdot \sqrt{1-x^2} \cdot exp(-\xi \cdot (1-x^2)) \qquad 0 < x < 1 \qquad (3.11)$$

where $x = m_{\rm ES}/m_{\rm ES}^0$, ξ is the parameter determining the slope and $m_{\rm ES}^0$ is the end-point of the distribution. We use the same function ($P_{bhh}(m_{\rm ES})$ of Eq. 4.3) to describe all the bhh backgrounds ($hh = \pi\pi, K\pi,$ and KK). ΔE is described using a second order polynomial ($P_{bhh}(\Delta E)$ of Eq. 4.3), whose parameters are floated in the fit. To parameterize the shape of the Fisher discriminant, we use the sum of two Gaussians.

In total, we have:

- 2 parameters for $m_{\rm ES}$
- 2 parameters for ΔE
- 5 parameters for $\mathcal{F}$

All the parameters except the end-point of $m_{\rm ES}$ are floated in the nominal fit. In order to perform tests on the fit, such as Toy Monte Carlo studies, we use as a reference the values obtained from $B^0 \to h^+h'^-$ analysis.

3.6.1.3 Fit Validation

We validate the fit to the hh category using samples of Monte Carlo events, coming from a full simulation or from toy experiments. In Table 3.12, we give the result of the fit on a sample of pure signal or background Monte Carlo events.

Sample	$N_{\pi\pi}$	$N_{K\pi}$	$N_{b\pi\pi}$	$N_{bK\pi}$	N_{bKK}
89282 $B \to \pi\pi$ events	88859 ± 299	245 ± 30	163 ± 19	11.9 ± 5.9	3.0 ± 2.3
86505 $B \to K\pi$ events	277 ± 38	86133 ± 295	12.6 ± 5.3	76 ± 13	5.2 ± 4.4
32439 $c\bar{c}$ events	-2.3 ± 3.8	-46.4 ± 8.1	11916 ± 114	8445 ± 103	12128 ± 115
55782 uds events	72 ± 24	10 ± 16	33662 ± 191	15118 ± 137	6921 ± 89

Table 3.12. Fit result on $B\bar{B}$ and $q\overline{q}$ Monte Carlo samples.

In order to quantify any bias in the fit, we also perform a set of toy Monte Carlo experiments in which we generate realistic samples (containing events from hadronic B decays, and $q\overline{q}$ background events) according to the likelihood shape of each component and we fit them. As in the nominal fit, yields, *CP* asymmetries and $q\overline{q}$ shape parameters are floated. As already specified, leptonic events are neglected in this case.

In order to check the fact that the fit is able to determine the shapes of background PDF's, we look for biases in the pull distributions of background shape parameters, shown in Figure 3.11. No significative bias is observed.

We estimate these biases by multiplying the means of the Gaussian fits to the pulls by the errors of the corresponding fits to the generated events, shown in Table 3.13. All the biases are small compared to the results of the fit; the biases for the $q\overline{q}$ yields are bigger than the others, but this is expected since it is a common feature to the $B^0 \to h^+h'^-$ analysis.

In addition, the pull distribution for the yields are shown in Figure 3.10.

3.6.1.4 Fit on Run1-5 hh on-resonance data

The output of the fit on full hh on-resonance dataset is given in Table 3.13. The values of the $q\overline{q}$ yields and shape parameters will be used in the nominal fit of leptonic yields. The uncertainties returned by the fit will be used to evaluate the systematic error associated to the knowledge of these parameters in the leptonic fit.

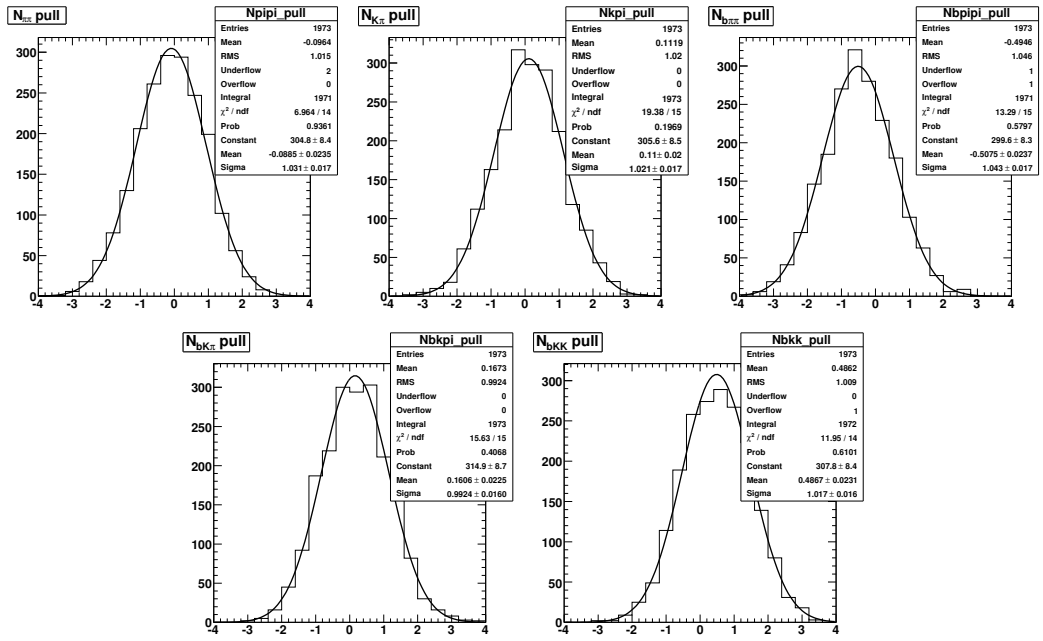

Figure 3.10. Distribution of the pull yields for the hadronic decays and $q\overline{q}$ components, obtained from a set of toy Monte Carlo experiments performed with the expected inputs on Run1-5 dataset for the hh PID category.

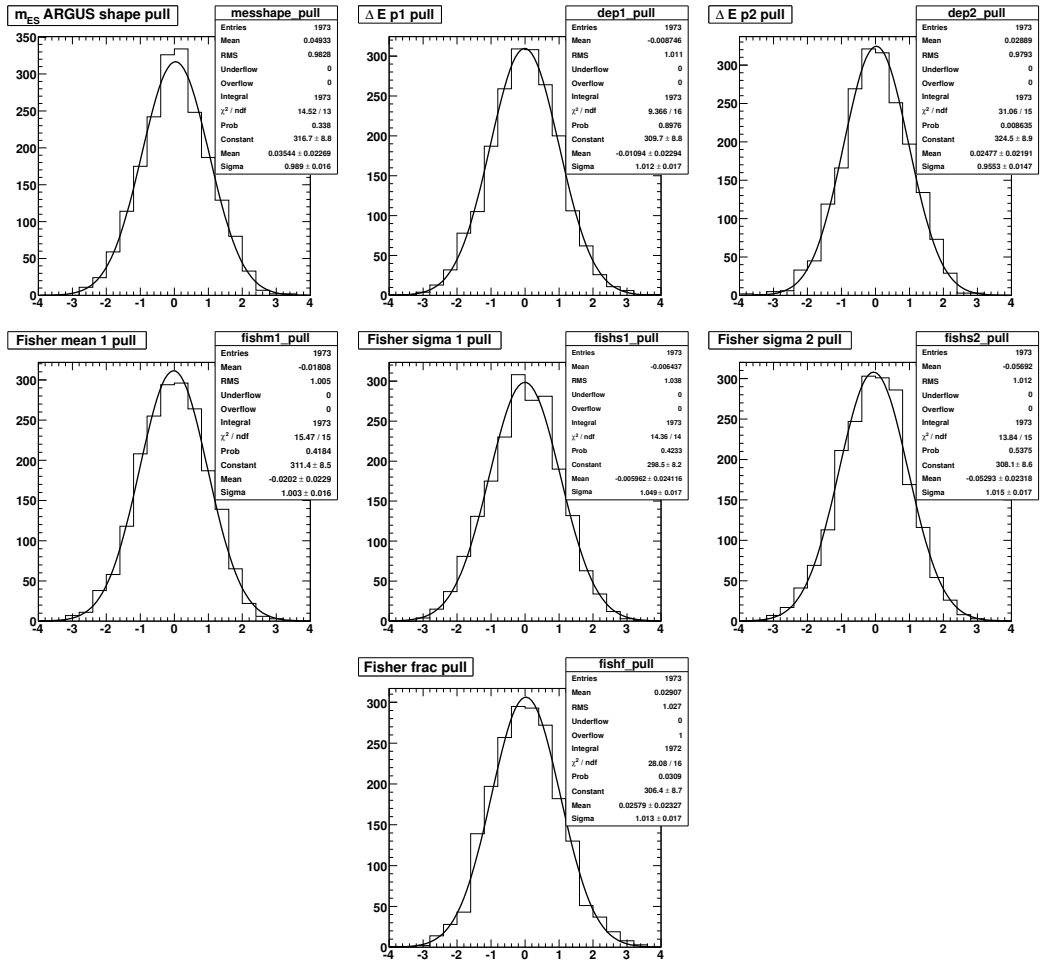

Figure 3.11. Distribution of the pull of the PDF shapes for the $q\overline{q}$ components, obtained from a set of toy Monte Carlo experiments performed with the expected inputs on Run1-5 dataset for the hh PID category.

Parameter	Result	Bias
$N_{\pi\pi}$	605 ± 39	-4
$N_{K\pi}$	2205 ± 59	7
$A_{CP}^{K\pi}$	-0.108 ± 0.026	0.001
$N_{b\pi\pi}$	35035 ± 199	-113
$N_{bK\pi}$	23401 ± 177	33
$A_{CP}^{bK\pi}$	-0.0033 ± 0.0075	-0.0000
N_{bKK}	15752 ± 135	74
SIG $m_{\rm ES}$ μ	5.27940 ± 0.00007	0.00000
SIG ΔE μ	-0.0016 ± 0.0018	-0.000098
BKG $m_{\rm ES}$ ξ	-19.17 ± 0.42	0.01
BKG ΔE p_1	-0.771 ± 0.065	0.000
BKG ΔE p_2	0.28 ± 0.77	-0.01
BKG $\mathcal{F}$ μ_1	0.3280 ± 0.0029	-0.000045
BKG $\mathcal{F}$ μ_2	0.052 ± 0.022	0.001
BKG $\mathcal{F}$ σ_1	0.3968 ± 0.0028	-0.000013
BKG $\mathcal{F}$ σ_2	0.639 ± 0.013	-0.001
BKG $\mathcal{F}$ f_1	0.883 ± 0.013	0.000

Table 3.13. Result of the fit performed on hh on-resonance Run1-5 data. The third column is the bias on the yields and on the PDF shapes estimated with 2000 toy Monte Carlo experiments. Each bias is calculated by multiplying the mean of the pull by the error of the fit of the toy Monte Carlo experiments. All the biases of the shape parameters of the PDF's are negligible, the only bias that will be used in the evaluation of the number of $q\overline{q}$ background events in the leptonic fits is the one to $N_{b\pi\pi}$.

3.6.2 Fit of Leptonic yields

The three leptonic yields are extracted from three independent fits, performed on each PID sample ($N_{\mu\mu}$ from 2μ events, $N_{e\mu}$ from $1\mu 1e$ events, and N_{ee} from $2e$ events). The structure of the fit is similar to that of Section 3.6.1, except for the fact that the θ_C pull is not used in this case.

In each of the fits, we assume the background shapes obtained from hh data (see Section 3.6.1). All the shape parameters are fixed in the fit, while the number of signal and continuum background events are floated. The likelihood is written as:

$$\begin{aligned}\mathcal{L} \quad = \quad & \frac{e^{-(N_{ll}+N^{ll}_{b\pi\pi})/N}}{\sqrt[N]{(N_{ll}+N^{ll}_{b\pi\pi})!}} \cdot \prod_{i=1}^{N} \Big\{ \\ N_{ll}\cdot \quad & P_{ll}(m_{ES})_i \cdot P_{ll}(\Delta E)_i \cdot P_{ll}(\mathcal{F})_i + \\ N^{ll}_{b\pi\pi}\cdot \quad & P_{bhh}(m_{ES})_i \cdot P_{bhh}(\Delta E)_i \cdot P_{bhh}(\mathcal{F})_i \Big\} \end{aligned} \tag{3.12}$$

where N_{ll} is the yield of the leptonic decay populating the considered PID class and $N^{ll}_{b\pi\pi}$ is the $q\overline{q}$ background in the same sample. The P functions of the background component are those used in the previous fits to hh data sample, while those for signal are obtained from unbinned maximum likelihood fits to samples of Monte Carlo events, corrected according to the study of appendix A.

3.6.2.1 Parameterization of signal PDF's

We parameterize leptonic signal PDF's using ML fits to signal Monte Carlo samples, after applying the PID weighting to improve the data/Monte Carlo agreement. These values (quoted in Table 3.14) are corrected according to the scaling factors of appendix A, in order to further improve the data/Monte Carlo agreement.

For $m_{\rm ES}$ we use different shapes for the different leptonic species, since the amount of radiation is different for electrons with respect to the other particles. This effect generates differences in the left-side tails of the kinematic variables. For $\mu\mu$, we use the same function used for hadronic events, since no difference is observed with

Variable	$\mu^+\mu^-$	$e^\pm\mu^\mp$	e^+e^-
SIG $m_{\rm ES}$ μ	5.27960 ± 0.00007	5.27970 ± 0.00007	5.28010 ± 0.00002
SIG $m_{\rm ES}$ α_L	0.1014 ± 0.0011	0.0914 ± 0.0015	0.0950 ± 0.0016
SIG $m_{\rm ES}$ σ_L	0.00253 ± 0.00002	0.00289 ± 0.00002	0.00310 ± 0.00002
SIG $m_{\rm ES}$ α_R	0.	0.0129 ± 0.0028	0.0195 ± 0.0034
SIG $m_{\rm ES}$ σ_R	0.00248 ± 0.00001	0.00248 ± 0.00001	0.00248 ± 0.00001
SIG ΔE μ	-0.00135 ± 0.00030	0.00159 ± 0.00038	0.00112 ± 0.00052
SIG ΔE α_L	0.1815 ± 0.0021	0.2972 ± 0.0035	0.0947 ± 0.0064
SIG ΔE σ_L	0.0230 ± 0.0002	0.0275 ± 0.0004	0.0341 ± 0.0006
SIG ΔE α_R	0.0781 ± 0.0026	0.0805 ± 0.0032	0.0947 ± 0.0040
SIG ΔE σ_R	0.0254 ± 0.0002	0.0255 ± 0.0003	0.3923 ± 0.0004
SIG $\mathcal{F}$ μ	-0.0327 ± 0.0041	-0.0327 ± 0.0041	-0.0327 ± 0.0041
SIG $\mathcal{F}$ σ_L	0.6840 ± 0.0028	0.6840 ± 0.0028	0.6840 ± 0.0028
SIG $\mathcal{F}$ σ_R	0.3835 ± 0.0025	0.3835 ± 0.0025	0.3835 ± 0.0025

Table 3.14. Summary of the fit parameters, obtained with 290k events of Monte Carlo of the three leptonic decays considered in the fits.

respect to the distribution of $B \to \pi\pi$ events. On the other hand, because of the tails induced by QED final state radiation, this is not true in the case of $e\mu$ and ee. The sets of parameters for the two $m_{\rm ES}$ functions are obtained from fits to signal Monte Carlo, without the constraint $\alpha_R = 0$, but forcing the resolution parameter σ_R of the right side tale to be the same as $P_{\pi\pi}(m_{ES})$ (since it is generated by detector resolution in tracks' reconstruction and it depends on the momentum more than on the nature of the particle). Figure 3.12 shows the distribution of the $m_{\rm ES}$ shapes for $B\bar{B}$ fit components, with the over-imposed fit.

As Figure 3.1 shows, the shape of ΔE strongly depends on the final state, since ΔE is strongly correlated to the energy of the radiated (and not reconstructed) photon. Even in this case the shape of $\mu\mu$ is close to the distribution of $B \to \pi\pi$ Monte Carlo, while the presence of electrons in the final state introduces deviations in the tails. We use three different Cruijff functions to describe the three shapes, as shown in Figure 3.13.

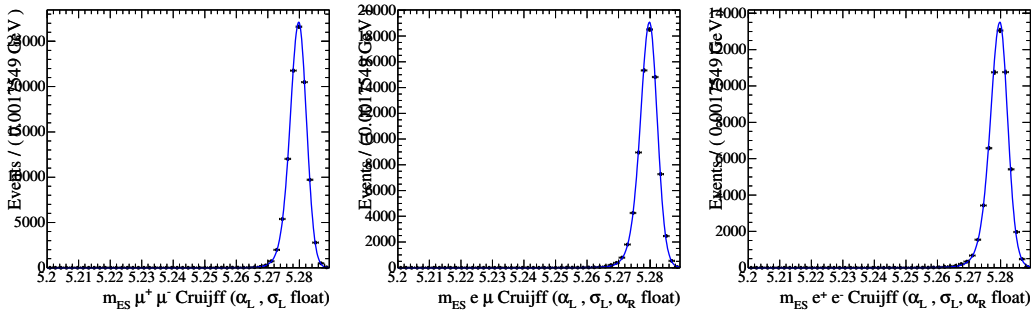

Figure 3.12. Fit to $m_{\rm ES}$ distributions for $\mu^+\mu^-$ (left), $e^\pm\mu^\mp$ (center), and e^+e^- (right), obtained from signal Monte Carlo samples.

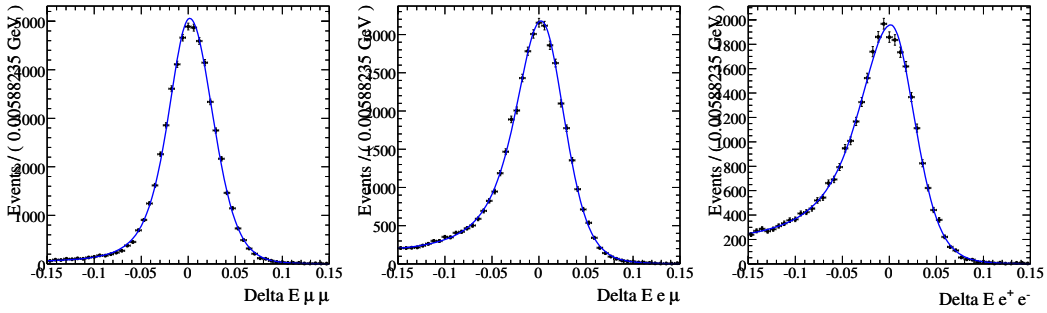

Figure 3.13. Fit to ΔE distributions for $\mu^+\mu^-$ (left), $e^\pm\mu^\mp$ (center), and e^+e^- (right), obtained fitting signal Monte Carlo samples.

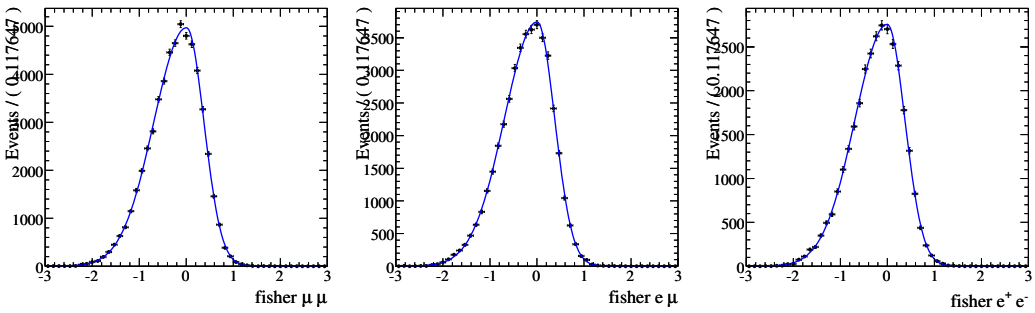

Figure 3.14. Fit to $\mathcal{F}$ distributions for $\mu^+\mu^-$ (left), $e^\pm\mu^\mp$ (center), and e^+e^- (right), obtained from signal Monte Carlo samples.

As for $B \to \pi\pi$ events, we described the Fisher discriminant using a bifurcated Gaussian. Since a sizable difference is observed, with respect to hadronic events, we use a different set of parameters in this case. On the other hand, the three leptonic shapes are consistent among each other, so that a common parametererization is used. Figure 3.14 shows the distribution for the three $B\bar{B}$ Monte Carlo samples.

In Table 3.15, we summarize the parameters used for the three sets of PDF's, after applying the data/Monte Carlo correction. All these parameters are fixed in the nominal fit and varied within their errors to evaluate the systematic errors. The errors, also quoted in the table, come from combining the errors from $B \to ll$ and $B \to \pi\pi$ Monte Carlo samples and the error from the fit on the hh data sample.

3.6.2.2 Fit Validation

We perform a set of toy Monte Carlo experiments separately for the three leptonic yields; each set is done by generating samples with the expected number of background events (as calculated in Tab. A.2), and the leptonic yields set to zero. These samples are then fitted keeping all the PDF shapes fixed, whereas the signal and background yields are floated in the fit. The pull distributions for the leptonic yields are shown in Figure 3.15. We cannot expect to fit these pulls with a Gaussian function because of Poissonian fluctuations induced by the small signal yield. In order to prove that

Variable	$\mu^+\mu^-$	$e^\pm\mu^\mp$	e^+e^-
SIG $m_{\rm ES}$ μ	5.27940 ± 0.00007	5.27950 ± 0.00010	5.27990 ± 0.00010
SIG $m_{\rm ES}$ σ_L	0.00227 ± 0.00009	0.00259 ± 0.00012	0.00277 ± 0.00013
SIG $m_{\rm ES}$ σ_R	0.00229 ± 0.00017	0.00229 ± 0.00017	0.00299 ± 0.00017
SIG $m_{\rm ES}$ α_L	0.189 ± 0.040	0.171 ± 0.020	0.177 ± 0.020
SIG $m_{\rm ES}$ α_R	0.041 ± 0.036	-0.087 ± 0.015	-0.131 ± 0.023
SIG ΔE μ	-0.0016 ± 0.0019	0.0013 ± 0.0018	0.0009 ± 0.0019
SIG ΔE σ_L	0.0267 ± 0.0008	0.0247 ± 0.0016	0.0305 ± 0.0017
SIG ΔE σ_R	0.0311 ± 0.0004	0.0313 ± 0.0011	0.482 ± 0.016
SIG ΔE α_L	0.112 ± 0.020	0.45 ± 0.18	0.144 ± 0.058
SIG ΔE α_R	0.003 ± 0.057	0.131 ± 0.012	0.145 ± 0.012
SIG $\mathcal{F}$ μ	-0.002 ± 0.039	-0.002 ± 0.039	-0.002 ± 0.039
SIG $\mathcal{F}$ σ_L	0.717 ± 0.022	0.717 ± 0.022	0.717 ± 0.022
SIG $\mathcal{F}$ σ_R	0.423 ± 0.030	0.423 ± 0.030	0.423 ± 0.030

Table 3.15. Summary of the fit parameters, obtained correcting the values of Table 3.14 according to the data/Monte Carlo comparison study of app. A.

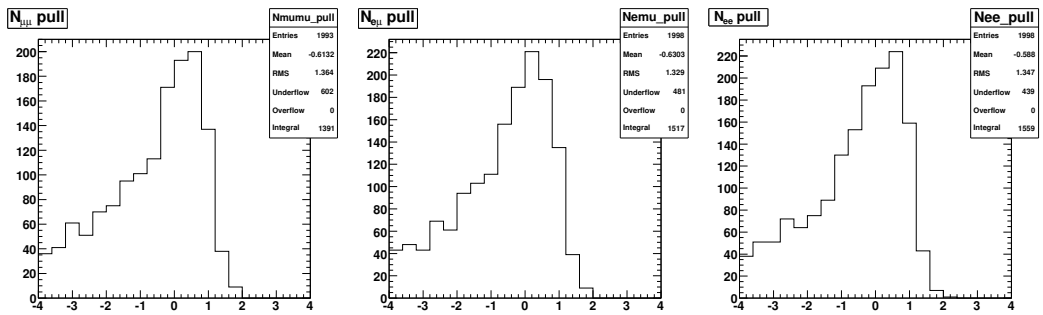

Figure 3.15. Distribution of the pull yields for the three signal decays from 2000 toy MC experiments, setting the mean value for the leptonic decays to zero.

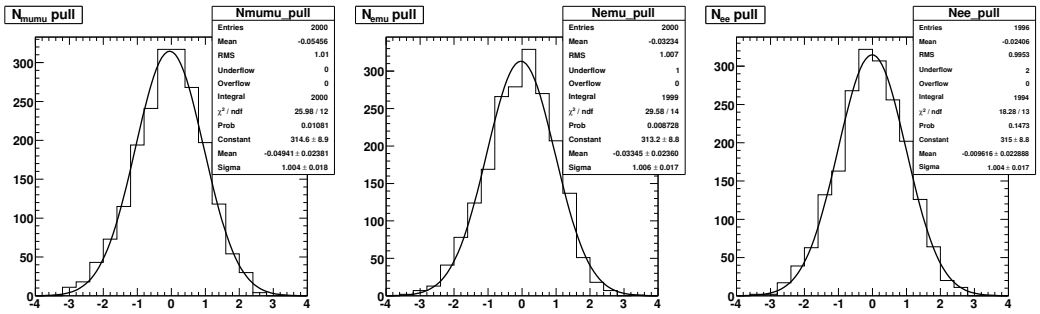

Figure 3.16. Distribution of the pull yields for the three signal decays from 2000 toy MC experiments, setting the mean value for the leptonic decays to one hundred and multiplying by one hundred the corresponding expected number of $q\overline{q}$ events.

the fits are unbiassed, besides this intrinsic non Gaussian effect, we generate samples with leptonic yields set to 100 and background yields one hundred times higher than expected whereas we keep all the other parameters fixed to the same value. The corresponding pulls are shown in Figure 3.16. From the pull distributions, we calculate the biases on the yields:

- $N_{\mu\mu}$: bias is -0.7 events;
- $N_{e\mu}$: bias is -0.5 events;
- N_{ee}: bias is -0.2 events;

which are negligible, when compared to 100 events we used in generation.

3.7 Systematics

The calculation of the Branching Ratio is affected by three different sources of systematic errors:

- systematic error on the yields;
- statistical and systematic error on the efficiencies;

- statistical and systematic error on the luminosity calculation.

In the following sections, we discuss each of these contributions.

3.7.1 Yields

The systematic error on the yields comes from the knowledge of the signal and background parameters. The background parameters are determined from a fit to the hh sample. The associated error is used to evaluate the systematic one, shifting each parameter by $\pm 1\sigma$ from its nominal value and quoting the difference on the leptonic yield as the error. A detailed list of the contributions is given in appendix A.

We use the same procedure for signal parameters, but in this case the central values and errors are obtained correcting the Monte Carlo parameters according to the study of appendix A, which takes into account the data/Monte Carlo agreement for the shapes of the PDF's (computed using the $B \to hh$ events of the hh sample). Since the errors on the parameters are calculated by propagating the errors on the single quantities (the value of the parameter from $B \to ll$ Monte Carlo, from $B \to \pi\pi$ Monte Carlo and from the fit to the hh sample) and since the dominant – which is the determination of the parameters from the fit to hh, limited by the available statistics – is a common source for all the parameters of the PDF, we quote as the total systematic error for each PDF the largest observed deviation from the nominal value. The breakdown of the contributions to the total error are given in appendix A.

The summary of the systematic errors on the yields of the leptonic decays is reported in Table 3.16.

3.7.2 Error on Efficiencies

The error on the knowledge of reconstruction efficiency is in part related to the available Monte Carlo statistics (binomial error on the efficiencies) and in part to the data/Monte Carlo agreement in the distribution of the variables used to select the final sample.

decay mode	$q\bar{q}$ shapes		Signal shapes		Total	
$N_{\mu\mu}$	$+0.065$	-0.070	$+0.380$	-0.450	$+0.385$	-0.455
$N_{e\mu}$	$+0.080$	-0.087	$+0.146$	-0.163	$+0.166$	-0.184
N_{ee}	$+0.532$	-0.532	$+1.065$	-0.930	$+1.190$	-1.056

Table 3.16. Summary of systematic errors on the number of signal events associated to the fit to the determination of signal and background shapes.

We consider and combine the selection efficiencies (estimated with Monte Carlo samples and reported in Table 3.2) with the efficiencies of the PID selectors, reported in Table 3.8.

We report in Table 3.17 a summary of these contributions, for each of the three leptonic decays considered. In particular:

- the systematic error on the selection efficiencies, taken from the $B^0 \to h^+h'^-$ analysis [56];
- the error on the efficiency related to the PID requirements (see Section 3.7.3).

Source	$\epsilon_{\mu\mu}$	$\epsilon_{e\mu}$	ϵ_{ee}
Standard Selection			
tracking	1.6	1.6	1.6
DIRC	1.0	1.0	1.0
$\cos\theta_S$	1.1	1.1	1.1
PID corrections	0.13	0.81	4.06
Total systematic	2.19	2.33	4.61
Statistical	0.19	0.22	0.30

Table 3.17. Summary of relative (%) statistical and systematic errors on efficiencies; "Standard Selection" refers to reconstruction and to the preliminary selection cuts described in Section 4.3; "PID corrections" refer to the error due to the use of the PID tables, described in Section 3.7.3; "Total systematic" is their sum in quadrature; "Statistical" is the total statistical error on the efficiencies, due to all the the selection criteria and to the statistical error on the PID tables. All the errors will be added in quadrature in the estimation of the upper limit on the Branching Ratio.

3.7.3 Systematic error on the efficiency due to the PID tables

The PID tables – which we use to correct for data / Monte Carlo disagreements – introduce errors in the fit, coming from the statistical and systematic error on the PID efficiencies. The tables are divided into bins of $\mathbf{p}$, θ and ϕ; for each bin, we know the efficiency and the statistical error on it, obtained from dedicated control samples. In addition, the value of the total efficiency, evaluated on the considered Monte Carlo sample, changes according to the algorithm used to apply the correction (on-line methods and off-line mode).

The estimation of this spread can be calculated using two different sets of PID tables: the ones based on data samples and the ones based on MC samples. For each event of the three Monte Carlo leptonic samples, we apply the following procedure:

1. for each bin of $\mathbf{p}$, θ and ϕ, we randomly generate a value of the efficiency according to a Gaussian distribution, centered in ε_i (the nominal value provided by the PID table for that bin of $\mathbf{p}$, θ and ϕ) and with RMS $\sigma(\varepsilon_i)$ (its associated error). This procedure is repeated both for Monte Carlo and data tables. In this way, we obtain alternative PID tables, distributed around the nominal ones according to the errors provided by the control samples.

2. we find the bin in the PID tables corresponding to each track of each Monte Carlo event which passes the PID requirement;

3. we calculate the ratio of the efficiency of the data-based table over that of the MC-based $r(\varepsilon_i)$ and its error $\sigma(r(\varepsilon_i))$;

The total efficiency ϵ is then calculated as $\epsilon = \frac{1}{N} \cdot \sum_{i=0}^{N} r(\varepsilon_i)^{(1)} \cdot r(\varepsilon_i)^{(2)}$ where N is the total number of events of the Monte Carlo sample, $r(\varepsilon_i)^{(1)}$ and $r(\varepsilon_i)^{(2)}$ are the efficiencies of the two track of the i-th event, only for those events that passed the PID requirements. We repeat this procedure 500 times for the three Monte Carlo samples, obtaining a distribution for the PID efficiencies. The RMS of this

distribution is the statistical errors on the efficiencies of the PID selectors, induced by the statistical uncertainty on the PID tables. Histograms for $\mu\mu$, $e\mu$ and ee are shown in Figure 3.17.

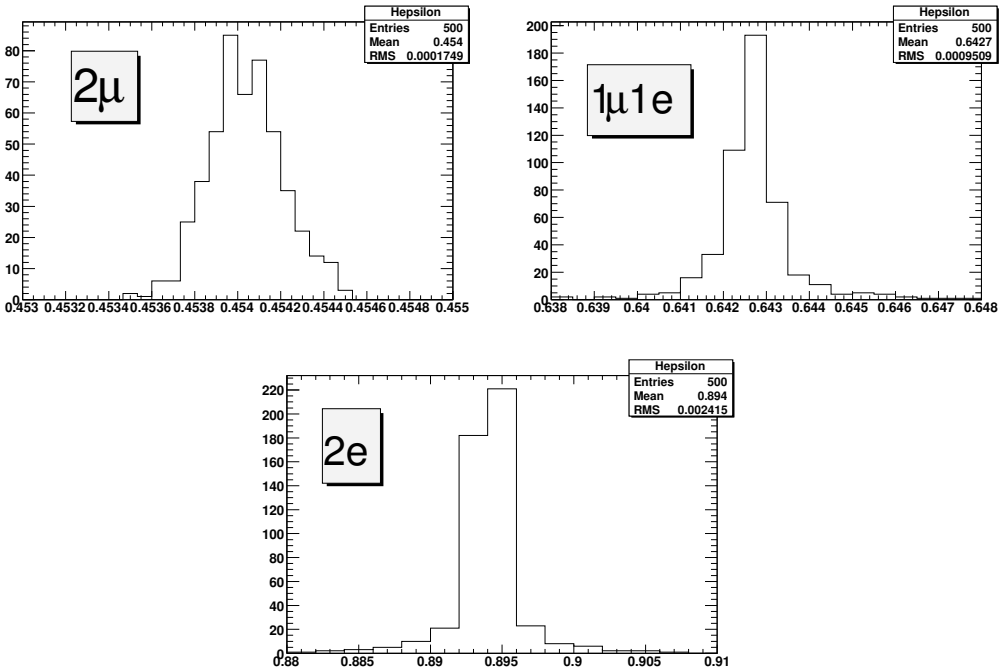

Figure 3.17. Distribution of the efficiencies of the PID tables, for $\mu\mu$, $e\mu$ and ee, obtained following the procedure described in the text. The RMS of these distributions quantifies the statistical error associated to the PID tables.

The systematic error for each PID class is given by the difference of the values obtained using the official on-line method and the mean of the distributions of Figure 3.17 (off-line method).

The efficiencies of the three classes, evaluated with the two different methods (by using the official PID tables and by using the means of the Gaussians of the off-line method) and without the PID correction are summarized in Table 3.18, whereas the corresponding systematic errors are summarized in Table 3.19.

The statistical error associated to the PID tables is usually smaller than the systematic one. The biggest discrepancy (about 4%) is observed in the $2e$ sample.

class	official PID tables	off-line method	no PID correction
2μ	45.272	45.404	50.037
$1\mu 1e$	65.082	64.270	65.815
$2e$	85.346	89.405	85.552

Table 3.18. Efficiencies (%) of the three leptonic classes, evaluated with two different methods: a) using the official PID tables, b) using the means of the Gaussians of the off-line method. The last column is the efficiency of the classes without the PID correction.

class	off-line method (syst.)	statistical
2μ	0.132	0.017
$1\mu 1e$	0.812	0.095
$2e$	4.059	0.241

Table 3.19. Summary of systematic error (%) and statistical error (%) for each of the three PID classes. The systematic error is the difference between the efficiency of the classes using official PID tables and the means of the Gaussians of the off-line method; the statistical error is the RMS of the Gaussians of the off-line method.

track	region	official PID tables	off-line method
μ^+	all	67.18 ± 0.15	67.210 ± 0.015
μ^+	FWD	68.25 ± 0.22	68.243 ± 0.020
μ^+	barrel	66.33 ± 0.20	66.082 ± 0.022
μ^+	BWD	50.00 ± 11.18	49.52 ± 1.75
μ^-	all	66.95 ± 0.15	66.983 ± 0.016
μ^-	FWD	67.45 ± 0.22	67.706 ± 0.025
μ^-	barrel	66.54 ± 0.20	66.40 ± 0.20 to check
μ^-	BWD	62.50 ± 12.10	72.261 ± 1.46
e^+	all	92.51 ± 0.11	94.86 ± 0.33
e^+	FWD	92.62 ± 0.17	96.90 ± 0.62
e^+	barrel	92.43 ± 0.15	93.19 ± 0.025 to check
e^+	BWD	1.00 ± 0.00	1.02 ± 1.75
e^-	all	92.43 ± 0.11	94.187 ± 0.025
e^-	FWD	93.05 ± 0.16	95.061 ± 0.030
e^-	barrel	91.93 ± 0.15	93.492 ± 0.036
e^-	BWD	1.00 ± 0.00	to do

Table 3.20. Efficiencies (%) of the three leptonic classes, evaluated with two different methods: a) using the official PID tables, b) using the means of the Gaussians of the off-line method. The last column is the efficiency of the classes without the PID correction.

3.7.4 Error on Number of $B\bar{B}$ events.

We take the official uncertainty on the number of $B\bar{B}$ events: $N_{B\bar{B}} = (382.9 \pm 4.2) \cdot 10^6$ events.

3.8 Evaluation of the Upper Limit

Assuming the SM predictions for the leptonic BR's, we do not expect any significant signal yield in the Run1-5 dataset. Since we know from the Section 3.6.2.2 that the ML fit does not provide an unbiassed estimator of signal yields near the physical boundary $N > 0$, because of Poissonian fluctuations (which invalidate the Gaussian assumption), we decided to use the entire shape of the likelihood function to quote the result, given in terms of an Upper Limit (UL). The UL is calculated using a flat prior for $N > 0$. Assuming Gaussian shapes of the systematic errors, we include the effect of the systematics in the UL evaluation. The entire procedure is explained below.

We assign Gaussian shapes to the systematic error on the yield, to the uncertainty on reconstruction efficiency and to the error on the number of $B\bar{B}$ pairs in the data sample. We calculate the BR likelihood using a toy Monte Carlo technique, which allows to perform numerically the convolution between the signal likelihood and the Gaussian smearings induced by the systematic errors. The BR likelihood is computed as the distribution of the quantity

$$\frac{N_{sig} + \sigma_{sys}(N_{sig})}{(\epsilon + \sigma(\epsilon)) \cdot (N_{B\bar{B}} + \sigma(N_{B\bar{B}}))} \tag{3.13}$$

where ϵ and $N_{B\bar{B}}$ are generated according to Gaussian shapes, N_{sig} is extracted according to its likelihood (which incorporates the statistical error), and $\sigma_{sys}(N_{sig})$ is the systematic shift, which is generated according to a Gaussian function, centered in zero and with RMS given by the systematic error.

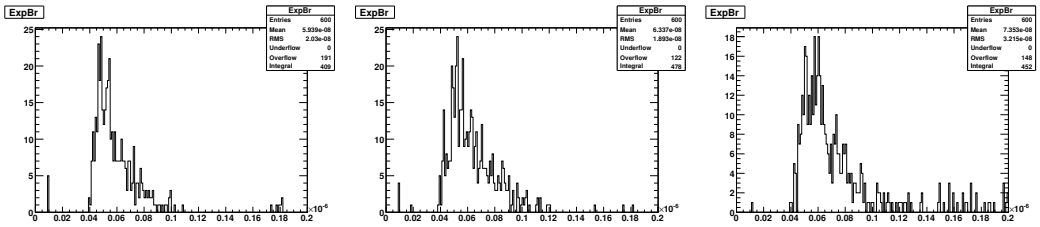

Figure 3.18. Distribution of the UL on $BR_{\mu\mu}$ (left) , $BR_{e\mu}$ (center), and BR_{ee} (right) from independent sets of toy Monte Carlo experiments..

decay	$UL(BR)(10^{-8})$
$B \to \mu\mu$	5.9
$B \to e\mu$	6.3
$B \to ee$	7.4

Table 3.21. Expected upper limits on the Branching Ratios for the three leptonic decays, using 600 independent toy Monte Carlo experiments for each decay. Systematic errors are included in these estimations.

In order to evaluate the sensitivity of the analysis, we apply this technique to a set of 600 samples, generated assuming a luminosity of 350 fb^{-1}, with the expected leptonic yields fixed to 0. We scan each of the three leptonic yields ($N_{\mu\mu}$, $N_{e\mu}$, and N_{ee}) in 100 steps in the range $[-5, 15]$. For each toy Monte Carlo sample, we assign to each scanned value of $N_{\mu\mu}$ the maximum of the likelihood (obtained fitting for the background yield). This provides us a shape for the signal yield, which is translated into a likelihood function for the BR, as explained above. Using this likelihood, we quote the 90% probability UL on the BR, the value of $UL(BR)$ such that

$$\int_0^{UL(BR)} dBR_{ll}\mathcal{L}(BR_{ll}) = 0.90 \int_0^{\infty} dBR_{ll}\mathcal{L}(BR_{ll}) \tag{3.14}$$

In this way, we obtain the distribution of UL values of the Branching Ratio given in Figure 3.18 and in Table 3.21.

3.9 Results

We perform separately the three leptonic fits on the full Run1-5 dataset. Figure 3.19 shows the likelihood distribution for the three leptonic decays, as function of the signal yields. The corresponding numbers of signal and background events, evaluated at the maxima of the likelihoods, are reported in Table 3.22. They are used only as cross-check, but are not used in the upper limit calculation.

The distributions of the likelihoods as function of the BR are shown in Figure 3.20, which also include all the systematic errors. We calculate the UL on the BR of the three leptonic decays by calculating the BR corresponding to the 90% of their area, as described in detail in Section 3.8. The UL on the branching fractions of the three leptonic decays are reported in Table 3.23.

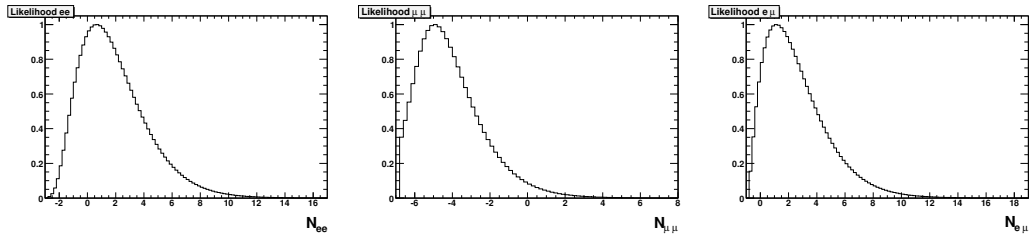

Figure 3.19. Distribution of the likelihood as function of the signal yield for e^+e^- (left), $\mu^+\mu^-$ (center) and $e^\pm\mu^\mp$ (right) decays.

Sample	Signal yield	Background yield
2μ	-4.94 ± 1.42	60.9 ± 8.2
$1\mu 1e$	1.13 ± 1.82	84.9 ± 9.3
$2e$	0.64 ± 2.08	66.4 ± 8.4

Table 3.22. Result of the fit on the three leptonic samples.

Our result on $B \to e^\pm\mu^\mp$ improves the previous *BABAR* result [57]. Because of the the different statistical approach followed, in the case of $B \to e^+e^-$ and $B \to \mu^+\mu^-$ the comparison should be made with the previous UL without background subtraction. Even in this case, we observe an improvement in the $\mu\mu$ decay. The better UL in the

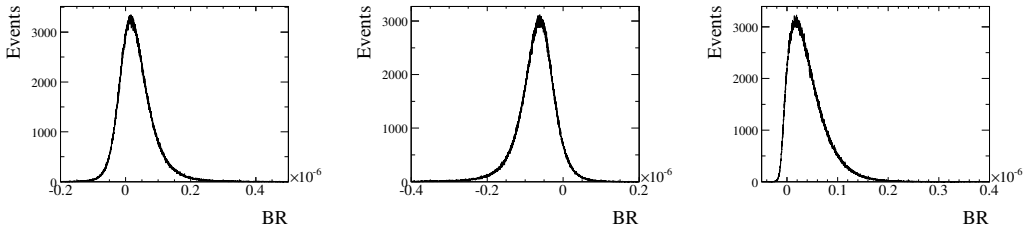

Figure 3.20. Distribution of the likelihoods as function of the BR for e^+e^- (left), $\mu^+\mu^-$ (center) and $e^\pm\mu^\mp$ (right) decays, which include all the systematic errors.

decay	$UL(BR)(10^{-8})$
$B \to \mu\mu$	5.2
$B \to e\mu$	9.2
$B \to ee$	11.3

Table 3.23. Upper upper limits on the Branching Ratios for the three leptonic decays. Systematic errors are included in these estimations.

e^+e^- channel quoted in [57] corresponds to the UL that this analysis would achieve if we integrated the experimental likelihood from $-\infty$ rather than from 0.

This work has been published on Phys. Rev. D **77**, 032007 (2008).

CHAPTER 4

CP ASYMMETRY IN $B \to X_S\gamma$ DECAYS

4.1 Introduction

There are several ways to measure the properties of $b \to s\gamma$ decays. From the experimental point of view, the best approach is to measure separately different exclusive decay channels to reduce the backgrounds. However, inclusive approaches are preferred from the theoretical point of view because they reduce the model dependence and the associated uncertainties. The work presented in this thesis is the analysis of $B \to X_s\gamma$ decays on the recoil of fully-reconstructed hadronic B's. It is an inclusive analysis because there is no attempt to characterize the structure of the hadronic X_s system.

A fully inclusive analysis is very complex and is studied elsewhere [61]. We reconstruct both B mesons in the $\Upsilon(4S) \to B\bar{B}$ decays. One of the two B's of the event, called $B_{\rm tag}$, is reconstructed if it decays in one among several very well known hadronic modes. The recoiling B, called $B_{\rm sig}$, is reconstructed if it decays into a photon of at least 1.3 GeV in the B rest frame. If there is more than one photon which satisfies this requirement, we choose the one with the highest momentum. The hadronic reconstruction of $B_{\rm tag}$ has several advantages:

- given the high backgrounds associated with the reconstruction of a single photon, this technique provides a much cleaner environment due to the separation of particles produced by the decay of the two B's in the event. This feature reduces the associated systematic uncertainties;

- the four-momentum of the $B_{\rm tag}$ is well measured. This enables us to measure the signal photon energy directly in the $B_{\rm sig}$ rest frame;
- we can tag the charge and flavor of the $B_{\rm tag}$, from which we can measure the direct *CP* asymmetry;
- continuum events can be estimated using the kinematics of the tagged B meson.

The disadvantage of this technique is the very limited statistics, because the hadronic decays of $B_{\rm tag}$ represent $\sim 5\%$ of the total B decays, which leads to an overall signal reconstruction efficiency of $\sim 0.5\%$

We identify four types of backgrounds: continuum events; combinatoric events, i.e. mis-reconstructed photons, mainly from the $B_{\rm tag}$ side of the event; photons coming from π^0 and η decays; background coming from other B decays. If the high energy photon is a decay product of a π^0 or η, we remove it by applying a cut on the invariant mass distribution of the corresponding two-photon system; continuum, combinatoric events and other B decay products are estimated in a maximum likelihood fit, which will extract the direct *CP* asymmetry.

The fit is built with two variables: $m_{\rm ES}$ and the output of a boosted decision tree (BDT) [58]. The quantity $m_{\rm ES}$ is calculated with the decay products of the $B_{\rm tag}$, whereas we use a selection of shape variables to build the BDT multivariate discriminant. We train the BDT in order to maximize the separation between signal and continuum events. There is a non-negligible component of B-background events in the peaking signal region: these events are estimated with MC simulation and fixed in the fit.

As already mentioned in Section 1.6, very little is known on the extrapolation uncertainty from higher to lower photon energies. The *CP* asymmetry is expected to be the same at any point of the photon energy spectrum, because $b \to s\gamma$ decays are determined by only one operator in the SM. In case more operators were needed,

their interference could lead to different values of CP asymmetries at different points of the photon spectrum. We select events with $E_\gamma > 2.2$ GeV: this is a very different cut value with respect to the previous analysis [62]. The choice is based on optimization studies aimed at minimizing the uncertainty on A_{CP}, performed with MC simulation.

4.2 Datasets

This analysis is based on data taken with the *BABAR* detector during Run 1 through 6, corresponding to $\sim 4.7 \times 10^8$ B meson pairs (equivalent to an integrated luminosity of 426 fb^{-1} at the $\Upsilon(4S)$ peak).

Table 4.1 shows information about all the simulated events used in this analysis: the number of generated events for each decay, the number of reconstructed hadronic tag B's and the assumed cross sections of each decay.

Decay mode	# Events	# $B_{\rm tag}$	σ(nb) or $\mathcal{B}$
$B^0 \to K^{*0}(892)\gamma$	6.1×10^6	32474	$(4.2 \pm 0.6) \times 10^{-5}$
$B^\pm \to K^{*\pm}(892)\gamma$	6.3×10^6	36562	$(4.2 \pm 0.6) \times 10^{-5}$
$B^\pm \to X_{su}\gamma$ (KN465)	20.6×10^6	225783	$(3.29 \pm 0.33) \times 10^{-4}$
$B^0 \to X_{sd}\gamma$ (KN465)	20.6×10^6	199378	$(3.29 \pm 0.33) \times 10^{-4}$
uds	9.4×10^8	1.9×10^7	2.09
$c\bar{c}$	1.1×10^9	3.9×10^7	1.30
B^+B^-	7.3×10^8	3.4×10^7	0.55
$B^0\bar{B}^0$	7.4×10^8	3.0×10^7	0.55

Table 4.1. MC samples used in this analysis. We report the decay mode, the number of generated events, the number of reconstructed hadronic B's and the assumed cross section or branching fraction.

4.2.1 Signal simulation

As already mentioned in Section 1.5, we use two kinds of signal MC for this analysis. The first one contains a non-resonant inclusive $b \to s\gamma$ decay model. The corresponding MC modes, indicated as KN465, are generated following the Kagan and

Neubert model [47] in which we set the b-quark mass to 4.65 GeV. Figure 4.1 shows the MC-true photon energy spectrum for the inclusive modes used in this analysis and the corresponding spectrum for the modes used in the previous analysis [62], in which the b-quark mass was set to 4.80 GeV. The current value of the b-quark mass is $m_b = 4.68^{+0.17}_{-0.07}$ GeV/c^2 [50].

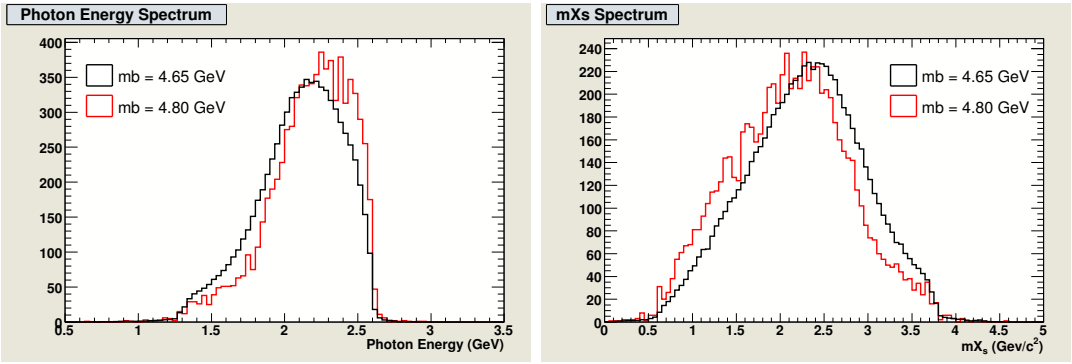

Figure 4.1. Distribution of the MC-true photon energy spectrum (left) and corresponding mass of the X_s system (right) for inclusive signal events generated with the quark mass of $m_b = 4.65$ GeV/c^2 (black curve) and $m_b = 4.80$ GeV/c^2 (red curve).

The second kind of signal MC contains only resonant exclusive $B \to K^{(*)}\gamma$ decays. These modes are characterized by the sharp $K^*(892)$ peak over the continuum of all the other resonances, which have very large widths and larger mass.

Figure 4.2 shows the MC-true energy spectrum of the inclusive and exclusive (resonant) signal MC samples. We combine the two samples according to the method suggested by Kagan and Neubert: we first scale the number of exclusive events to the same integrated luminosity of the inclusive ones, preserving their branching fractions ratio; then we define a cut-off value of the X_s mass, $m_{X_s}^{\text{cut-off}} = 1.15$ MeV, and we combine events from the inclusive sample with MC-true $m_{X_s} > m_{X_s}^{\text{cut-off}}$ with events from the exclusive sample with MC-true $m_{X_s} < m_{X_s}^{\text{cut-off}}$. The resulting MC-true photon energy spectrum is shown in Figure 4.3. The choice of $m_{X_s}^{\text{cut-off}}$ will lead to a systematic error taken into account in later sections.

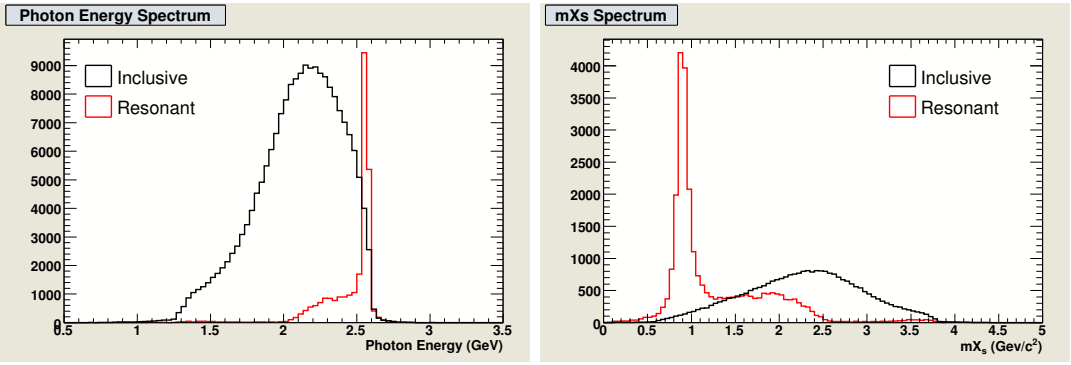

Figure 4.2. Distribution of the MC-true photon energy spectrum (left) and corresponding mass of the X_s system (right) for inclusive signal events (black curve) and resonant $K^{(*)}\gamma$ events (red curve).

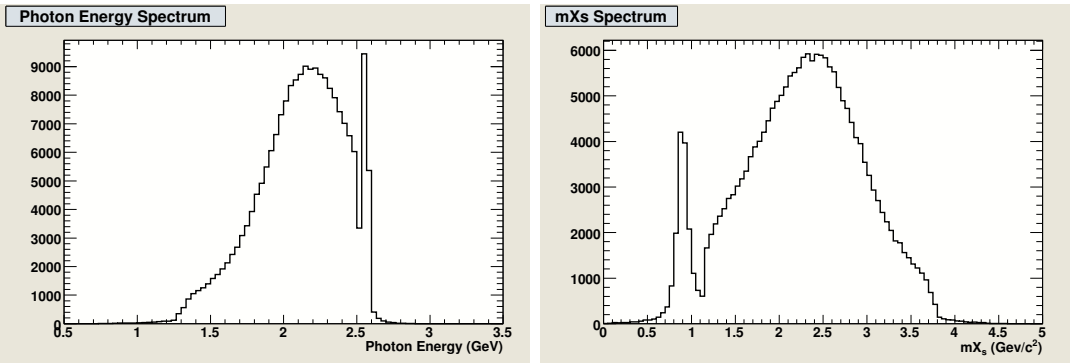

Figure 4.3. Distribution of the MC-true photon energy spectrum (left) and corresponding mass of the X_s system (right) for the combination of inclusive and resonant $K^{(*)}\gamma$ events.

4.2.2 Background simulation

We simulate two kinds of background: continuum jet-like non-$B\bar{B}$ events and $B\bar{B}$ events. The first type is composed of $\sim 1.3 \times 10^9$ generated $c\overline{c}$ events and $\sim 9.4 \times 10^8$ generated $u\overline{u}$ or $d\overline{d}$ or $s\overline{s}$ events, mixed according to their relative cross sections, reported in Table 4.2. We normalize the $c\overline{c}$ events to the corresponding

$e^+e^- \rightarrow$	Cross section (nb)
$b\bar{b}$	1.05
$c\bar{c}$	1.30
$u\bar{u}$	1.39
$d\bar{d}$	0.35
$s\bar{s}$	0.35

Table 4.2. Cross sections of $b\bar{b}$ and continuum events produced at the center of mass energy $\sqrt{s} = m(\Upsilon(4S))$.

integrated luminosity of the other type of continuum events. This results in a total integrated equivalent luminosity of 449 fb^{-1}. This is the sample with the lowest available statistics.

The generic $B\bar{B}$ sample is composed of a total of 1398 fb^{-1}. It is the full simulation of all possible decays of the B meson and it should be an unbiased event sample. When studying the $B\bar{B}$ background, signal events must be removed from this sample. Generic $B\bar{B}$ events are identified as signal if the reconstructed high energy photon, the X_s system and the decaying B are truth-matched at the same time. Each generated particle in the event contains information about its nature; truth-matching is an algorithm that associates each reconstructed particle to the corresponding generated nature.

4.3 Reconstruction and event selection

In this section we describe the event selection and optimization procedure. We begin by describing the reconstruction of the hadronic B, followed by the selection of

tracks, of the high energy photon and of all the other neutral clusters in the signal-B side; then we describe the procedure to discriminate signal photons that originate from π^0 and η decays. The training and application of a multivariate discriminant and the description of the maximum likelihood fit are illustrated in sections 4.4 and 4.5, respectively.

4.3.1 Hadronic reconstruction

The reconstruction of B decays into hadronic modes (*Semi-exclusive* reconstruction) is a novel technique, well described in many *BABAR* documents (see [62]).

The aim of the Semi-exclusive reconstruction is to get as many as possible B mesons reconstructed in fully hadronic modes in order to study the properties of the recoiling B.

Since B^0 mesons mostly decay into charged $D^{(*)}$ mesons while B^- mesons decay into the neutral $D^{0(*)}$ mesons, only such modes are considered. Table 4.3 shows the relevant branching fractions of the B mesons decaying predominantly into fully hadronic final states.

The Semi-exclusive reconstruction comprises the following steps:

- reconstruct all possible decay modes $B \to DX$, where the X system is a combination of π^+,π^0, K^+ and K^0_S, with total charge equal to ± 1, and including a maximum of 7 particles, 5 charged tracks and 2 neutrals;

- study the structure of the X system looking for resonances in the signal and studying the shape of the background (Section 4.3.1.2);

- identify submodes and create subcategories according to the their multiplicity and to the structure of the X system (e.g. $D\pi\pi^0$, $M_{\pi\pi^0} < 1.5\,\mathrm{GeV}/c^2$). For each mode, the most relevant parameter is the *apriori-purity* of the mode: the ratio $S/\sqrt{S+B}$ where S and B are the signal and combinatorial background respectively, as estimated from an m_{ES} fit on data (Section 4.3.1.2);

mode	branching fraction (%)
$B \to D^{*\pm} X$	22.5 ± 1.5
$B \to D^{\pm} X$	23.9 ± 1.9
$B \to D^{*0}/\overline{D}^{*0} X$	26.0 ± 2.7
$B \to D^{0}/\overline{D}^{0} X$	63.9 ± 3.0
$B \to D_s^{+} X$	10.5 ± 2.6
$B \to D^{-(*)} D_s^{(*)}$	4.8 ± 1.2
$B \to D^{(*)} \overline{D}^{(*)} K$	7.1 ± 2.3
$B^0 \to D^{-(*)} D^{+(*)}$	~ 1.0

Table 4.3. Inclusive and Exclusive branching fractions relevant to this analysis as measured in [50].

- determine a mode by mode combinatorial background rejection, in order to account for different background levels depending on the number of charged tracks and, above all, on the number of π^0 s in the reconstructed mode (Section 4.3.1.3);
- rank the submodes according to their purity and yields and study the significance as a function of the number of used modes in order to maximize the statistical significance of the sample (Section 4.3.1.4);
- group the submodes with similar purity;
- resolve the multiple candidates (Section 4.3.1.4).

The starting point of the Semi-exclusive selection is the D^0, D^+, D^*, D^{*0} meson reconstruction as described in Sec. 4.3.1.1.

Next, clean lists of charged pions and kaons, π^0, and K_S^0 are needed to be combined to the D meson to form the B candidated.

Pairs of opposite charge hadrons ($V^0 = h^+h^-$) and quartets of hadrons ($W^0 = h^+h^-h^+h^-$) are created using the list of charged tracks. If both the K_S^0 decay products are among the tracks used for a V^0 or a W^0, the two tracks are replaced by the K_S^0 (i.e. a V^0 would become a K_S^0 and a W^0 either a K_S^0 $\pi\pi$ or a K_S^0 K_S^0).

We accept B candidates with $m_{\rm ES} > 5.28\,{\rm GeV}/c^2$ and in a ΔE windows varying from 30 MeV to 80 MeV depending on the mode.

4.3.1.1 D reconstruction

The D mesons is reconstructed in a large variety of channels listed in Table 4.4.

Decay mode	**B.F.(%)**
$D^* \to D^0\pi$; $D^0 \to K\pi$	2.55 ± 0.06
$D^* \to D^0\pi$; $D^0 \to K3\pi$	5.0 ± 0.2
$D^* \to D^0\pi$; $D^0 \to K\pi\pi^0$	8.8 ± 0.6
$D^* \to D^0\pi$; $D^0 \to K^0_S\pi\pi$ ($K^0_S \to \pi^+\pi^-$)	1.35 ± 0.08
$D^+ \to K\pi\pi$	9.1 ± 0.6
$D^+ \to K^0_S\pi$ ($K^0_S \to \pi^+\pi^-$)	0.94 ± 0.06
$D^+ \to K\pi\pi\pi^0$	6.4 ± 1.1
$D^+ \to K^0_S\pi\pi\pi$ ($K^0_S \to \pi^+\pi^-$)	2.38 ± 0.31
$D^+ \to K^0_S\pi\pi^0$ ($K^0_S \to \pi^+\pi^-$)	3.5 ± 1.0
$D^{*0} \to D^0\pi^0$; $D^0 \to K\pi$	2.35 ± 0.12
$D^{*0} \to D^0\pi^0$; $D^0 \to K3\pi$	4.6 ± 0.3
$D^{*0} \to D^0\pi^0$; $D^0 \to K\pi\pi^0$	8.1 ± 0.7
$D^{*0} \to D^0\pi^0$; $D^0 \to K^0_S\pi\pi$ ($K^0_S \to \pi^+\pi^-$)	1.2 ± 0.1
$D^{*0} \to D^0\gamma$; $D^0 \to K\pi$	1.44 ± 0.19
$D^{*0} \to D^0\gamma$; $D^0 \to K3\pi$	2.82 ± 0.18
$D^{*0} \to D^0\gamma$; $D^0 \to K\pi\pi^0$	5.0 ± 0.4
$D^{*0} \to D^0\gamma$; $D^0 \to K^0_S\pi\pi$ ($K^0_S \to \pi^+\pi^-$)	0.7 ± 0.1
$D^0 \to K\pi$	3.80 ± 0.09
$D^0 \to K3\pi$	7.46 ± 0.31
$D^0 \to K\pi\pi^0$	14.0 ± 0.9
$D^0 \to K^0_S\pi\pi$	2.03 ± 0.12

Table 4.4. D Meson decay modes and the corresponding Branching Fractions as in [50].

The D^0 is reconstructed in the modes $D^0 \to K\pi$, $D^0 \to K3\pi$,$D^0 \to K\pi\pi^0$and $D^0 \to K^0_S\pi\pi$. The charged tracks originating from a D meson are required to have a minimum momentum of 200 MeV/c for the $D^0 \to K\pi$ and 150 MeV/c for the remaining three modes. The D^0 candidates are required to lie within $\pm 3\sigma$, calculated on an event-by-event basis, of the nominal D^0 mass. All D^0 candidates must have momentum greater than 1.3 GeV/c and lower than 2.5 GeV/c in the $\Upsilon(4S)$ frame. The

	$D^0 \to K\pi$	$D^0 \to K\pi\pi^0$	$D^0 \to K3\pi$	$D^0 \to K^0_S\pi\pi$
m_D invariant mass window	$\pm$ 15 MeV	$\pm$ 25 MeV	$\pm$ 15 MeV	$\pm$ 20 MeV
Charged tracks: lower $p*$ cut	$> 200\,\mathrm{MeV}/c$	$> 150\,\mathrm{MeV}/c$		
D^0 upper $p*cut$	$< 2.5\,\mathrm{GeV}/c$			
D^0 lower $p*cut$	$> 1.3\,\mathrm{GeV}/c$			
Vertex fit	$\chi^2 > 0.01$			

Table 4.5. Selection criteria for D^0 modes.

	$D^+ \to K\pi\pi$ $D^+ \to K^0_S\pi$	$D^+ \to K^0_S\pi\pi^0$	$D^+ \to K\pi\pi\pi^0$ $D^+ \to K\pi\pi\pi$
m_D invariant mass window	$\pm$ 20 MeV	$\pm$ 30 MeV	$\pm$ 30 MeV
D^+: lower $p*$ cut	$> 1.0\,\mathrm{GeV}/c$		$> 1.6\,\mathrm{GeV}/c$
D^+: upper $p*$ cut	$< 2.5\,\mathrm{GeV}/c$		
Charged tracks: lower p cut	$> 200\,\mathrm{MeV}/c$		
Vertex fit	$\chi^2 > 0.01$		

Table 4.6. Selection criteria for D^+ reconstruction.

lower cut is done to reduce combinatorics, the upper is due to the kinematic endpoint of the D^0 coming from a $B \to D^0X$ decay or $B \to D^{*+}X$ with $D^{*+} \to D^0\pi^+$. A vertex fit is performed, where a χ^2 probability greater than 0.1% is required. Selection criteria are summarized in Table 4.5.

D^+ candidates are reconstructed in the modes $D^+ \to K^-\pi^+\pi^+$, $D^+ \to K^-\pi^+\pi^+\pi^0$, $D^+ \to K^0_S\pi^+$, $D^+ \to K^0_S\pi^+\pi^0$ and $D^+ \to K^0_S\pi^+\pi^+\pi^+$. The minimum charged track momentum is required to be 200 MeV/c. D^+ candidates are required to have an invariant mass within $\pm 3\sigma$, calculated on an event-by-event basis, of the nominal D^+ mass. The D^+ candidates must have momentum greater than 1.0 GeV/c in the $\Upsilon(4S)$ frame for the three cleanest modes ($D^+ \to K^-\pi^+\pi^+$, $D^+ \to K^0_S\pi^+$ and $D^+ \to K^0_S\pi^+\pi^0$) and greater than 1.6 GeV/c for the two remaining ones ($D^+ \to K^-\pi^+\pi^+\pi^0$ and $D^+ \to K^0_S\pi^+\pi^+\pi^+$). Moreover all D^+ candidates must have momentum lower than 2.5 GeV/c in the $\Upsilon(4S)$ frame, as the D^0 case. A vertex fit is performed and a

Criteria	**Cut**
$D^{*+} \to D^0\pi^+$	
Vertexing and χ^2	beam spot constraint($\sigma_y = 30$ μm), convergence
$m(D^0\pi^+) - m(D^0)$	$\pm 3\sigma$ MeV/c^2
$p^*(\pi^+)$	[70,450] MeV/c
$D^{*0} \to D^0\pi^0$	
$m(D^0\pi^0) - m(D^0)$	± 4 MeV/c^2
$p^*(\pi^0)$	[70,450] MeV/c
$p^*(D^{*0})$	$1.3 < p^* < 2.5$ GeV/c
$D^{*0} \to D^0\gamma$	
$m(D^0\gamma) - m(D^0)$	[127,157] MeV/c^2
$E^*(\gamma)$	[100,450] MeV
$p^*(D^{*0})$	$1.3 < p^* < 2.5$ GeV/c

Table 4.7. Summary of cuts for D^{*+} and D^{*0} selection

χ^2 probability greater than 0.1% is required. Selection criteria are summarized in Table 4.6.

D^{*+} candidates are formed by combining a D^0 with a pion which has momentum greater than 70 MeV/c. Only the channel $D^{*+} \to D^0\pi^+$ is reconstructed since $D^{*+} \to D^+\pi^0$ events enter in the $B \to D^+X$ category. A vertex fit for the D^{*+} is performed using the constraint of the beam spot to improve the angular resolution for the soft pion. A fixed $\sigma = 30$ μm is used to model the beam spot spread in the vertical direction. The fit is required to converge, but no cut is applied on the probability of χ^2. After fitting, selected D^{*+} candidates are required to have Δm within $\pm 3\sigma$ of the measured nominal value.

D^{*0} candidates are reconstructed by combining a selected D^0 with either a π^0 or a photon having momentum less than 450 MeV/c in the $\Upsilon(4S)$ frame. The minimum momentum for the π^0 corresponds to 70 MeV while the photons are required to have an energy greater than 100 MeV. For $D^{*0} \to D^0\pi^0$, selected D^{*0} candidates are required to have Δm within 4 MeV/c^2 of the nominal value while the window is wider for $D^{*0} \to D^0\gamma$ ($127\,\mathrm{MeV}/c^2 < \Delta m < 157\,\mathrm{MeV}/c^2$).

Selection criteria for D^{*0} and D^{*+} are summarized in Table 4.7.

4.3.1.2 Study of the X system

A detailed study of the X system, looking for resonances in the signal and background shape was performed. This is meant to optimize the overall purity since a relatively dirty mode could perhaps be split into a very clean and a very dirty one. An example is shown in Figure 4.4, where the mode $B \to D\pi\pi\pi$ is analyzed.

In the upper plot, the a_1 line shape is clearly visible, but there is a significant contribution at higher masses. There is a large contribution above 1.5 MeV (non-resonant contribution and π_2). There is also a narrow structure around the D_s^+ mass which might be due to a non negligible $D_s^+ \to 3\pi$. Above 2 GeV/c^2 just a small amount of signal is present but the combinatoric background is very large, especially for the dirty D meson modes.

In order to further investigate this interpretation the lower plots of Figure 4.4 show the invariant masses plots of pairs of tracks in the X system for the a_1 ($M_X < 1.5\,\mathrm{GeV}/c^2$) and the π_2 ($1.5\,\mathrm{GeV}/c^2 < M_X < 2.\,\mathrm{GeV}/c^2$) regions separately. While the a_1 plot clearly shows a ρ signal, the π_2 shows both the f_2 and the ρ as expected from the decays of the π_2. To properly understand the final state the Dalitz analysis should be done, but this is not the purpose here that is meant to isolate dirty regions from clean ones. Two sub-modes are defined depending on whether M_X is smaller than 1.5 GeV/c^2 or it is between 1.5-2.0 GeV/c^2, without requiring the the sub-mode belonging to a precise resonance structure.

Finally the number of B modes is 52 (53 for the D^+ seed). The total number of modes is 1097. A summary is shown in Table 4.8.

4.3.1.3 ΔE selection

Once all the possible reconstruction modes are identified, a window in ΔE and a criterion to pick up among several candidates in a given mode have to be determined.

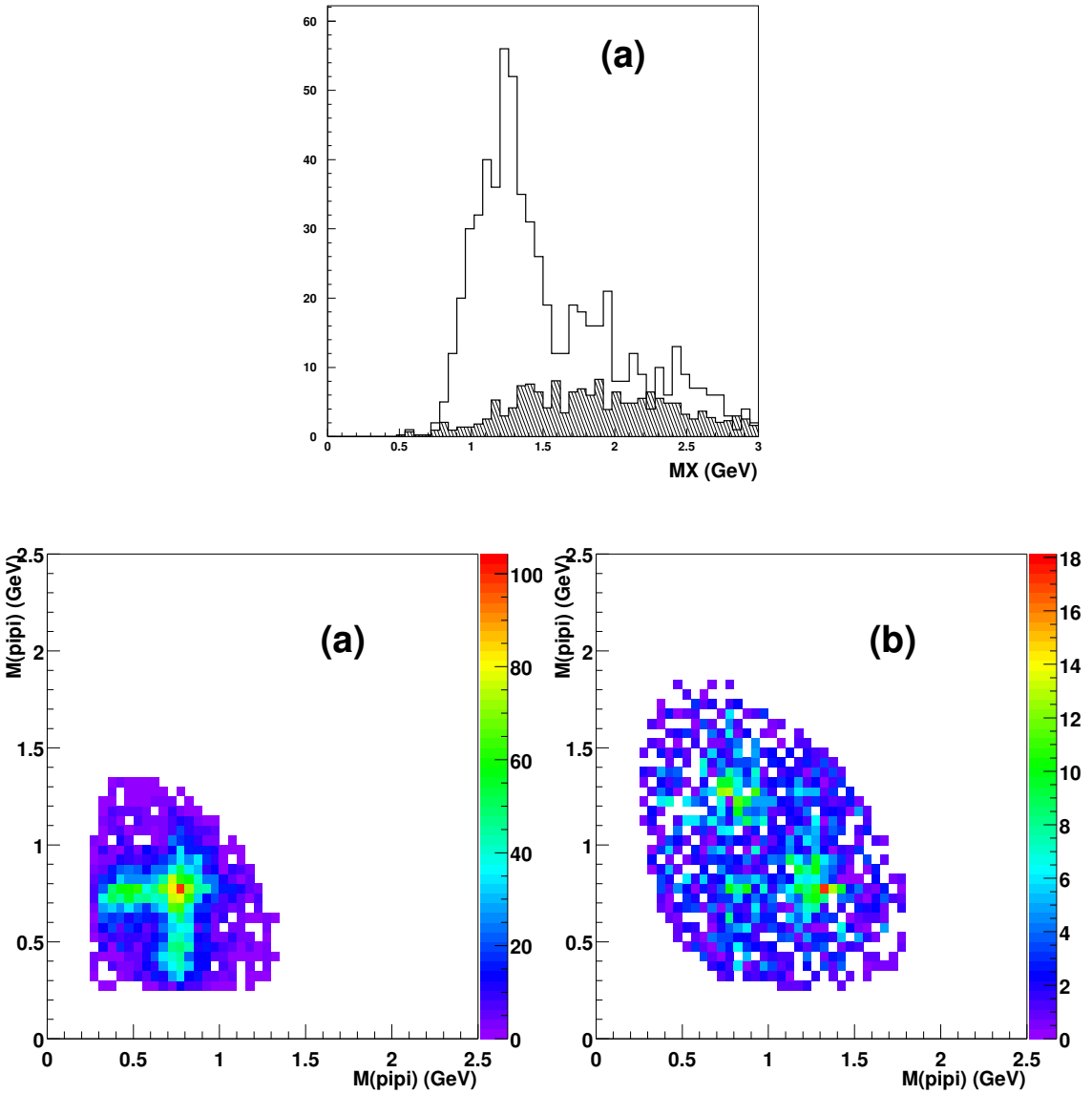

Figure 4.4. a) M_X distribution for the $D^*3\pi$ on the reduced sample (20 fb^{-1}). Only $D^* \to D^0, D^0 \to K\pi$ is plotted. The properly normalized background (hatched histogram), as evaluated from sidebands, is also shown. $M_{12} - M_{13}$ scatter plots for the three pions system for the mass region around the b) a_1 ($M_X < 1.5\,\mathrm{GeV}/c^2$) or c) the π_2 ($1.6 < M_X < 2.0\,\mathrm{GeV}/c^2$).

Channel	**pre-seed mode**	**# of B modes**	**total # of modes**
$B^+ \to D^0 X$	$D^0 \to K^-\pi^+$	52	208
	$D^0 \to K^-\pi^+\pi^0$	52	
	$D^0 \to K^0_S\pi^+\pi^-$	52	
	$D^0 \to K^-\pi^+\pi^+\pi^-$	52	
$B^0 \to D^+ X$	$D^+ \to K^-\pi^+\pi^+$	53	265
	$D^+ \to K^-\pi^+\pi^+\pi^0$	53	
	$D^+ \to K^0_S\pi^+$	53	
	$D^+ \to K^0_S\pi^+\pi^0$	53	
	$D^+ \to K^0_S\pi^+$	53	
$B^+ \to D^{*0} X$	$D^{*0} \to D^0\pi^0$, $D^0 \to K^-\pi^+$	52	416
	$D^{*0} \to D^0\pi^0$, $D^0 \to K^-\pi^+\pi^0$	52	
	$D^{*0} \to D^0\pi^0$, $D^0 \to K^0_S\pi^+\pi^-$	52	
	$D^{*0} \to D^0\pi^0$, $D^0 \to K^-\pi^+\pi^+\pi^-$	52	
	$D^{*0} \to D^0\gamma$, $D^0 \to K^-\pi^+$	52	
	$D^{*0} \to D^0\gamma$, $D^0 \to K^-\pi^+\pi^0$	52	
	$D^{*0} \to D^0\gamma$, $D^0 \to K^0_S\pi^+\pi^-$	52	
	$D^{*0} \to D^0\gamma$, $D^0 \to K^-\pi^+\pi^+\pi^-$	52	
$B^0 \to D^{*+} X$	$D^{*+} \to D^0\pi^+$, $D^0 \to K^-\pi^+$	52	208
	$D^{*+} \to D^0\pi^+$, $D^0 \to K^-\pi^+\pi^0$	52	
	$D^{*+} \to D^0\pi^+$, $D^0 \to K^0_S\pi^+\pi^-$	52	
	$D^{*+} \to D^0\pi^+$, $D^0 \to K^-\pi^+\pi^+\pi^-$	52	
TOTAL			1097

Table 4.8. Summary of the number of Semi-exclusive modes.

The ΔE resolutions are determined from the ΔE distributions before requesting the best candidates and they depend essentially on the number of charged tracks and, above all, on the number of π^0's in the only X system (since the reconstructed D meson is mass constrained). For the modes without π^0's a fit with a linear background and a Gaussian is performed and 2 σ symmetric windows are taken. In the case of modes with at least a π^0, the situation is worse. First of all there are too many candidates per event. Requiring that only the 10 candidates with the smallest $|\Delta E|$ are taken, can create a bias in the ΔE distribution. Therefore only the cleanest modes are used to determine a common window for all modes including π^0's. Moreover the presence of π^0 makes the distribution asymmetric.

The ΔE window varies from $-45 < \Delta E < 30$ MeV in the modes without π^0's, to $-90 < \Delta E < 60$ in the modes with 2 π^0's.

4.3.1.4 Multiple candidates and definition of purity

Two kinds of multiple candidates are possible: multiple candidates can be reconstructed in the same submode and many reconstructed submodes per event are also possible.

If there are multiple candidates in the same submode the candidate with the minimum ΔE is used and one candidate per submode is selected.

The selection of the best B among different sub-modes cannot use the ΔE criterion because the modes with higher combinatoric background would be privileged with respect to the clean ones, thus introducing a bias. The idea is to find an unbiased criterion for choosing a signal event based on a $a-priori$ probability. The $a-priori$ probability here is given by the purity of the mode, determined by fitting the $m_{\rm ES}$ distribution.

The selection of the best B in the event is based on the choice of the reconstructed mode with the highest purity.

The modes are ranked according to their purity and are added up to the sample one at a time. At each addition of a mode the yield increases and the purity decreases. This method is very useful once the composition of the modes has to be optimized for the analysis of the recoil. The significance $S/\sqrt{S+B}$ is computed as a function of the number of added modes and the best composition is chosen. An example for $B^0 \to D^{*+}X$ case is shown in Figure 4.5.

4.3.2 Track selection in the X_s system

The track selection is identical to that of the previous analysis [62] and is outlined below. It is summarized in Table 4.9.

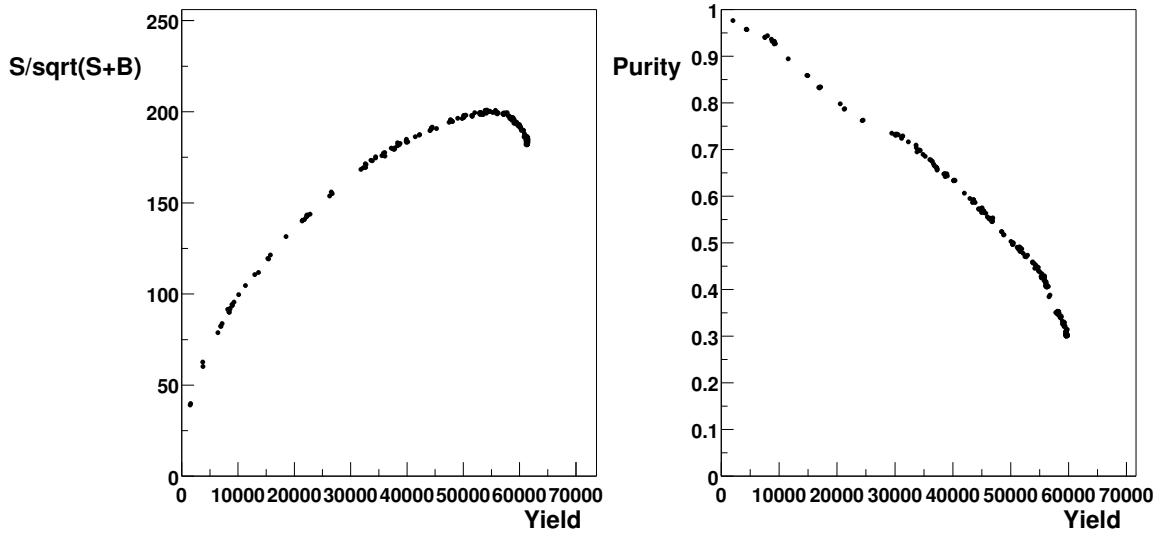

Figure 4.5. Dependence of the quality factor $S/\sqrt{S+B}$ as a function of the yield when adding modes for the $B^0 \to D^{*+}X$ case. Statistics corresponds to 80 fb^{-1}.

- The distance of closest approach to the beam spot must be less than 1.5 cm in the $x-y$ plane and less than 10 cm along the z axis. This removes tracks that do not originate from close to the interaction point. For secondary tracks from K_s^0 decays these restrictions are not imposed.

- For tracks with $p_\perp > 0.2$ GeV at least one DCH hit is required to remove poorly reconstructed tracks. This cut is not used for low momentum tracks since slow pions (for instance in the $D^* \to D^0\pi$ decays) would be rejected.

- A cut on the maximum momentum of $p_{\rm lab} < 10$ GeV, where $p_{\rm lab}$ is the laboratory momentum of the track, is applied. This removes tracks not compatible with the beam energies.

- Tracks are required to be within the polar angle acceptance of the detector: $0.410 < \theta_{\rm lab} < 2.54$ rad ensuring a well-understood tracking efficiency.

Select tracks with	Selection criteria
distance in $x-y$ plane	$\lvert d_{xy}\rvert < 1.5$ cm
distance in z axis	$\lvert d_z\rvert < 10$ cm
minimum number of DCH hits	$N_{\rm DCH} > 0$ if $p_\perp > 0.2$ GeV
maximum momentum	$p_{\rm lab} < 10$ GeV
geometrical acceptance	$0.410 < \theta_{\rm lab} < 2.54$ rad

Table 4.9. Summary of track selection criteria.

4.3.3 Neutral selection in the X_s system

The following criteria are applied to select "good" reconstructed photons from the X_s part of the signal-B, with the goal of removing photons from bremsstrahlung, fakes from hadronic interactions and unmatched clusters from charged particles. The selection requirements are summarized in Table 4.10.

- To ensure that the energy deposit is fully contained within the EMC, only clusters with polar angle $0.410 < \theta_{\rm lab} < 2.409$ rad are considered.

- A minimum energy, $E_\gamma^{lab} > 80\,\mathrm{MeV}$, is required to remove low energy photons associated with beam related backgrounds.

- The lateral moment (LAT) of a shower is a useful variable for rejecting background from hadronic interactions in the EMC (notably from K_L's or neutrons) that fake photons. LAT is defined as:

$$LAT = \frac{\sum_{i=3}^{N} E_i r_i^2}{\sum_{i=3}^{N} E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2} \quad , \tag{4.1}$$

  where N is the number of crystals within the cluster and E_i are the crystal energies in descending order, r_i is the distance of crystal i to the centre of the cluster and r_0 is the average distance between two crystals (approximately 5 cm for the *BABAR* calorimeter).

The value of LAT is defined to be in the range 0 to 1. LAT is small for electromagnetic showers where most of the shower's energy is deposited within a few crystals, and higher for hadronic showers where the energy deposit is less localized. This difference is enhanced by the two crystals with the highest energy deposits being omitted in the numerator, and by multiplying by squared distances from the shower cluster centre. LAT is required to be in the range $0.05 < LAT < 0.5$ for good photons.

- Another variable which discriminates against hadronic interactions is the ratio of the energy deposited in the 9 crystals closest to the cluster centroid, to the energy deposited in the 25 closest crystals ($S9/S25$). For electromagnetic showers $S9/S25$ is close to 1, whereas $S9/S25$ is smaller for hadronic showers as they deposit their energy in a larger number of crystals. For good photons $S9/S25$ is required to be greater than 0.9.

- Inefficiencies can occur in the track-cluster matching algorithm. If cluster energy is found close to a charged track but not matched to it, its deposit may be mistaken for a neutral cluster. If this occurs, the energy deposited is double counted. These clusters are removed by placing requirements on the distances in ϕ (dϕ) and θ (dθ) on the EMC surface to all tracks without a matched cluster.

4.3.4 High energy photon selection

We inherited the initial set of event selection criteria for signal photon candidates from the previous analysis [62], which consists of the following:

Photon energy

We select photons with energy $E_\gamma > 1.3$ GeV in the signal-B frame. The single photon energy spectrum is steeply rising at lower energies, which is due mainly to background decays. If there is more than one photon candidate from the signal-B

Description	Selection criteria
geometrical acceptance	$0.410 < \theta_{\rm lab} < 2.409\,{\rm rad}$
neutral energy	$E_\gamma > 80\,{\rm MeV}$
LAT	0.05 <LAT< 0.5
$S9/S25$	$S9/S25 > 0.9$
unmatched cluster	*Reject clusters if:* $\lvert d\theta\rvert < 30$ mrad, and $-30 < d\phi < 70$ mrad (to positive track) $-70 < d\phi < 30$ mrad (to negative track)

Table 4.10. Summary of the neutral selection criteria.

with energy exceeding 1.3 GeV, the most energetic photon is selected. The signal region is defined for photons with $E_\gamma > 2.2$ GeV: this choice of minimum energy is explained in section 4.5.

EMC cluster criteria

The EMC cluster associated with a photon must consist of at least 4 crystals to ensure that the photon is well reconstructed. We also require that the cluster is fully contained within the EMC ($0.410 < \theta_{lab} < 2.409$ rad).

Unmatched clusters

Cluster energy deposits found close to a track but not matched to it are removed. The criteria to reject unmatched clusters are detailed in Table 4.11. These requirements are the same as those applied to non-signal photons from the signal-B as described in section 4.3.3.

Lateral moment

LAT of a cluster is defined in Section 4.3.3. To reduce background from hadronic interactions in the EMC that fake a photon signature, LAT is required to be less than 0.50 for signal photon candidates. This requirement also helps to suppress background from merged π^0's, where the energy deposits from the photon decay products are so

close that they are indistinguishable as separate clusters. The requirement LAT$<$ 0.5 is chosen to avoid cutting in the region where the agreement between data and simulation is poor.

Photon isolation

Signal events have a lower multiplicity than continuum and B background events, so requiring that the selected photon is isolated from other energy deposits in the EMC helps discriminate against background. The discriminating variable is the distance between the photon cluster and the nearest bump in the EMC stemming from any other particle. A minimum distance of 25 cm on the EMC surface is required.

Description	Selection criteria
photon energy	$E_\gamma > 2.2$ GeV
geometrical acceptance	$0.410 < \theta_{\rm lab} < 2.409$ rad
LAT	$LAT < 0.5$
photon isolation	min bump distance > 25 cm
unmatched cluster	*Reject clusters if:*
	$\lvert d\theta \rvert < 30$ mrad, and
	$-30 < d\phi < 70$ mrad (to positive track)
	$-70 < d\phi < 30$ mrad (to negative track)

Table 4.11. Summary of the high energy photon selection criteria.

4.3.5 Selection optimization

The selection optimization consists of consecutive steps. The first step is the $B_{\rm tag}$ selection, already discussed in section 4.3.1. The remaining background is composed of three different contributions:

- photons coming from $\pi^0 \to \gamma\gamma$ or $\eta \to \gamma\gamma$ events;
- continuum $q\overline{q}$ events;
- other B decays.

The first of the three backgrounds is removed by choosing two discriminating variables and cutting on them, as we discuss in the next section; continuum and other B decays are estimated and subtracted with a maximum likelihood fit, as explained in sections 4.4 and 4.5.

4.3.6 π^0 and η invariant masses

The principal source of background comes from $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ decays. This background can be suppressed by eliminating photon candidates that are part of a photon pair with an invariant mass close to the nominal π^0 or η mass.

To identify π^0 decays, photon candidates are paired with all other photons in the X_s system and the pair with invariant mass closer to the π^0 mass is selected. Since the first photon is highly energetic, the second is restricted by kinematics to have low energy. The minimum cluster energy requirement of 80 MeV for "good" reconstructed photons is too high in this case. We simultaneously optimize the cut on the invariant mass distribution of the two-photon system and the minimum energy of the second photon. We scan a range of values for the two cuts and choose the best by maximizing the figure of merit (FOM) $S/\sqrt{S+B}$. In this case, S is simulated, truth-matched, signal and B is the combination of continuum and generic $B\bar{B}$ events in the peaking region, $m_{\rm ES} > 5.27$ GeV/c^2. All MC samples are scaled to an integrated luminosity of ~ 450 fb^{-1}.

The same procedure for reconstructing π^0 candidates can be applied to reconstructing η's. However, the η reconstruction is not as effective due to its higher mass. While for signal events it is rare to find a second photon that gives a mass as low as the π^0 mass, one almost always finds a second photon to get an invariant mass close to the η mass. In this case, it is necessary to require a higher energy for the second photon in order not to lose too much signal. Figure 4.6 shows the invariant mass distribution for reconstructed $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ events in all components.

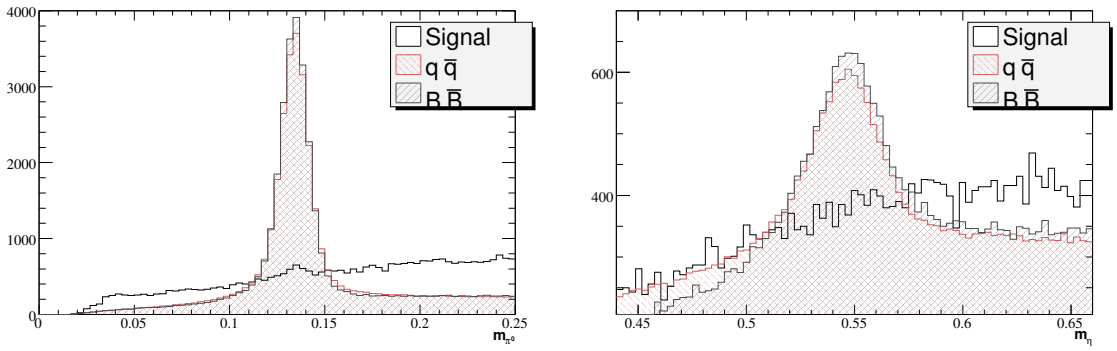

Figure 4.6. Invariant mass distribution of the reconstructed photon pair for $\pi^0 \to \gamma\gamma$ (left) and $\eta \to \gamma\gamma$ (right) decays.

The optimization of the two selection cuts is performed iteratively, *i.e.* we first find the highest value of FOM for the invariant mass, then we fix that value and optimize the minimum energy of the second photon. We fix the value thus obtained and scan the first variable again in a smaller range and so on. Figures 4.7 and 4.8 show FOM as function of the invariant mass distribution and of the minimum energy of the second photon for the π^0 and η cases, respectively. We notice a clear, albeit weak maximum in the FOM with respect to the two invariant mass distributions: they are the best chosen values. The maximum FOM also corresponds to the lowest values of the cut on the the minimum energy of the second photon. As already mentioned, this value is too small in the η case, therefore we keep the same cut as the previous analysis. Our optimal selection cuts are shown in Table 4.12.

We attempt to optimize a cut on the invariant mass distribution of a photon pair produced by a π^0 coming from ρ decays (notably, $B \to D^*\rho,\ \rho \to \pi^0\pi^+,\ \pi^0 \to \gamma\gamma$). The invariant mass distribution is shown in Figure 4.9, but we do not use this variable because the corresponding FOM does not show a maximum.

Decay mode	selection
$\pi^0 \to \gamma\gamma$	$m_{\pi^0} < 77$ MeV/c^2 or
	$m_{\pi^0} > 193$ MeV/c^2
	$E_{\gamma 2} > 30$ MeV
$\eta \to \gamma\gamma$	$m_\eta < 508$ MeV/c^2 or
	$m_\eta > 588$ MeV/c^2
	$E_{\gamma 2} > 120$ MeV

Table 4.12. Optimal cut selections for $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ events.

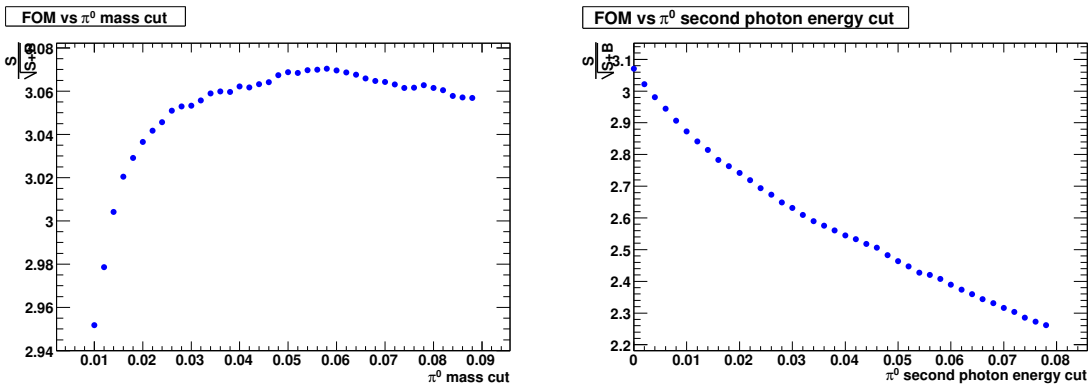

Figure 4.7. Optimization of the FOM vs the photon pair invariant mass (left) and vs the second photon minimum energy (right) for $\pi^0 \to \gamma\gamma$ decays. The horizontal axis on the left is the minimum value of the cut around the nominal π^0 mass, expressed in GeV/c^2. The zero on the horizontal axis on the right corresponds to a minimum second photon energy of 30 MeV. This scale is expressed in GeV.

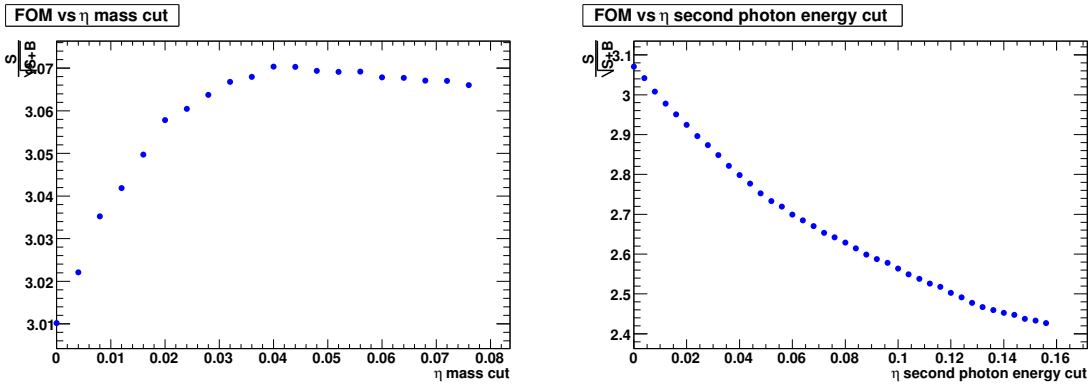

Figure 4.8. Optimization of the FOM vs the photon pair invariant mass (left) and vs the second photon minimum energy (right) for $\eta \to \gamma\gamma$ decays. The horizontal axis on the left is the minimum value of the cut around the nominal π^0 mass, expressed in GeV/c^2. The zero on the horizontal axis on the right corresponds to a minimum second photon energy of 30 MeV. This scale is expressed in GeV.

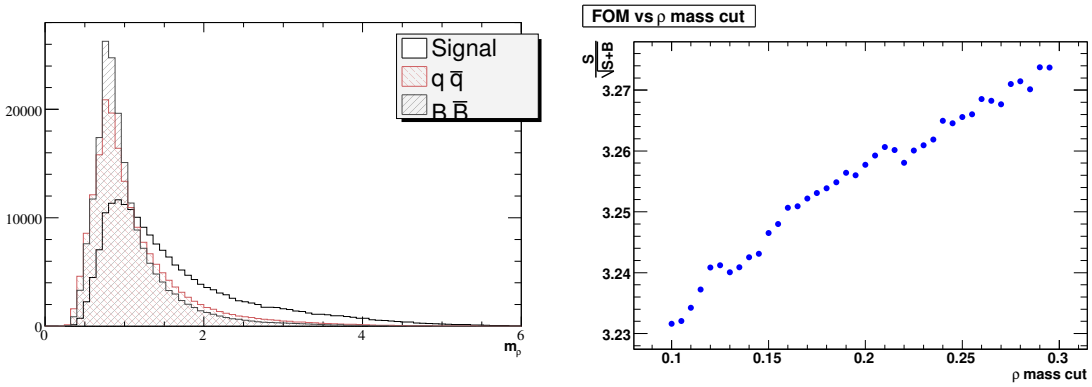

Figure 4.9. Invariant mass distribution of a reconstructed photon pair for π^0's produced in ρ decays (left) and optimization of the FOM (right). The horizontal axis on the right is the minimum value of the cut around the nominal ρ mass, expressed in GeV/c^2.

4.4 Multivariate discriminant

The first steps of event selections involve $B_{\rm tag}$ selection and removal of events from $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ decays. The remaining background is composed mainly of continuum events and of other generic B decays. Continuum events are typically estimated and subtracted by looking at the different topology with respect to signal (or, more in general, B-type events). It is relatively easy to construct variables able to discriminate between jet-like $q\overline{q}$ events and B decays, which present a homogeneous spatial distribution, since the B's are produced nearly at rest.

However, many discriminating variables present correlations among them, due to the fact that they are defined mostly using geometrical event distributions. The best way to take into account such correlations is to combine them into one single multivariate analysis (MVA) discriminant. This procedure has the ulterior advantage that a multivariate tool usually enhances the discriminating power with respect to using all the input variables separately. In order to reduce the statistical error on the *CP* measurement due to background, we include the MVA variable in a maximum likelihood fit.

Among the different possibilities, we choose to study three different variables: a Fisher discriminant [59], a neural network and a boosted decision tree. The choice of one over the others is based on two factors: the discriminating power between signal and background and the possibility to determine a functional shape of its distribution, in order to easily include it in the fit. The selection variables which we use to train the MVA discriminant are effective to separate signal from continuum events, because they are all related to event topology. For this reason, on the other hand, they are not effective to remove events from other B decays, since they are topologically very similar to our signal. This leads us to the choice of using only signal and $q\overline{q}$ events to train the MVA variable, but not generic $B\bar{B}$ events, which have to be estimated in a different way. Furthermore, we do not have a sufficiently large generic $B\bar{B}$ MC

sample for this purpose, especially if we request to use events with a stringent cut on the photon energy.

We use TMVA (A Toolkit for Multivariate Analysis), standalone version 3.9.5, to train our MVA discriminant and to build the corresponding distribution on all our MC samples.

4.4.1 Input variables

In this section, we describe all the variables used to train the MVA discriminant. All the distributions shown in this section assume a cut on the photon energy $E_\gamma > 2.1$ GeV, a minimum bump distance of 25 cm and are free from $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ events.

- **Thrust**

  The thrust axis of an event, $\hat{T}$, is the direction which maximizes the sum of the longitudinal momenta $\mathbf{p}$ of the particles. Thrust T, is related to $\hat{T}$ by

  $$T = \frac{\sum_i |\hat{T} \cdot \mathbf{p}_i|}{\sum_i |\mathbf{p}_i|}. \tag{4.2}$$

  T ranges from 0.5 for an isotropic event to 1 for a highly directional event. For continuum events the decay products of the B candidate lie in one of the two back to back jets, hence the decay axis of the B candidate is approximately collinear with the thrust axis of the rest of the event. For true B decays, the decay axis of the B candidate and thrust axis of the rest of the event are uncorrelated. We use the following thrust variables, shown in Figure 4.10:

  - the thrust computed with all the final particles in the event (T);
  - the thrust computed with only the decay products of the $B_{\rm tag}$ ($T_{B_{\rm tag}}$);
  - the absolute cosine of the angle between the thrust axis of $B_{\rm tag}$ daughters and the thrust of axis of $B_{\rm sig}$ daughters ($|\cos(\theta_{TBB})|$).

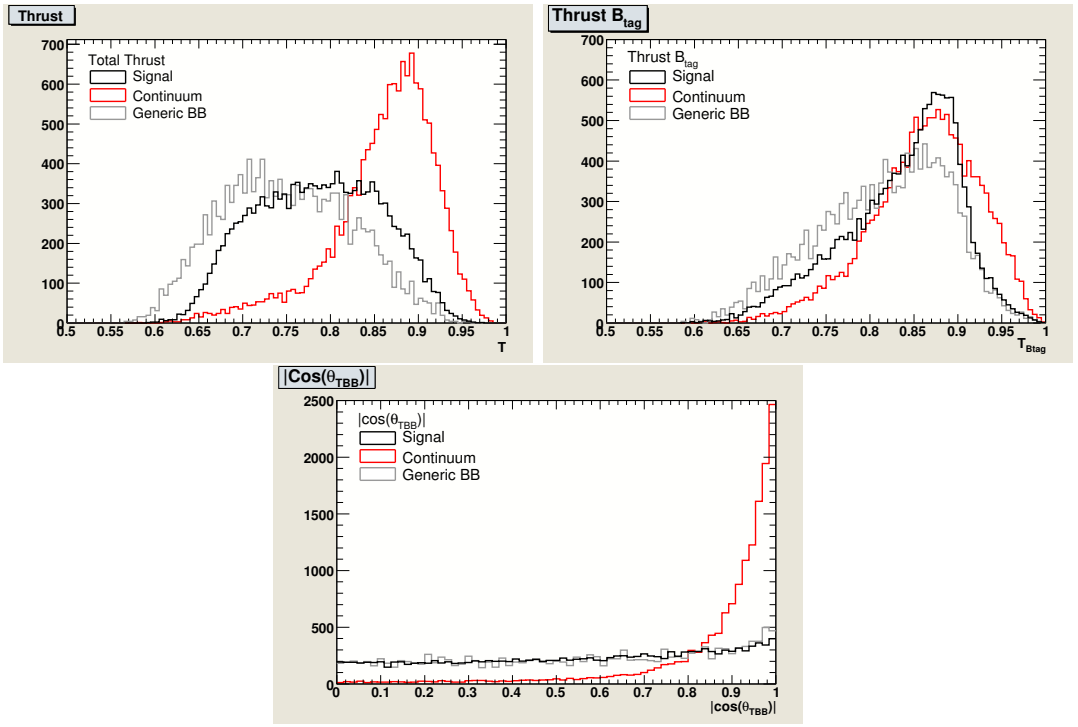

Figure 4.10. Thrust obtained with all particles in the event (left) and with only daughters of $B_{\rm tag}$ (right); cosine of the angle between the thrust axis calculated with $B_{\rm sig}$ daughters and the thrust axis calculated with $B_{\rm tag}$ daughters (bottom) for signal, continuum and generic $B\bar{B}$ events.

- **$\mathbf{R_2}$ and Sphericity**

  The definitions of these variables are given in Section 4.3. The corresponding distributions computed with all the particles in the event are shown in Figure 4.11.

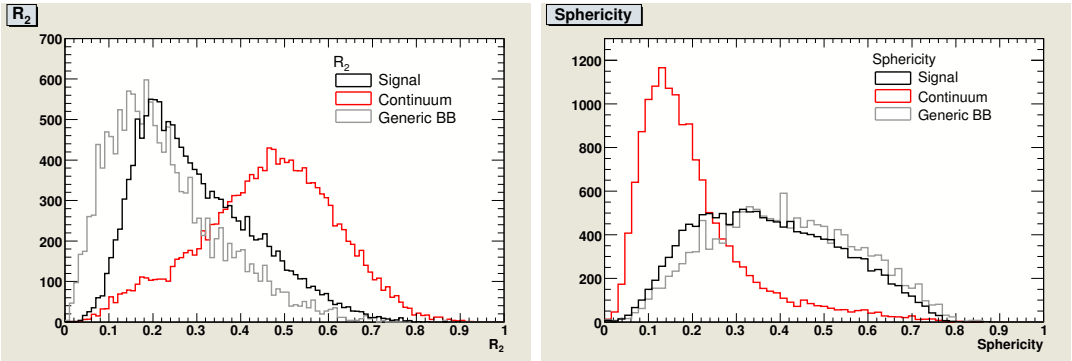

Figure 4.11. R_2 (left) and Sphericity (right) distributions for signal, continuum and generic $B\bar{B}$ events.

- **B_{tag} angle**

  The direction of a B meson from a $\Upsilon(4S)$ decay has a $\sin^2(\theta_B)$ angular distribution with respect to the beam axis in the $\Upsilon(4S)$ frame due to angular momentum conservation. The distribution of the cosine of the B_{tag} angle $|\cos(\theta_B)|$ is shown in Figure 4.12. For continuum events this distribution is expected to be flat.

- **Minimum bump distance**

  The cut of 25 cm on the minimum bump distance, described in section 4.3.4, ensures the reconstruction of good photons, i.e. not stemming from detector background events. After the cut, this variable is included in the training of the MVA discriminant and is shown in Figure 4.12.

- **B_{tag} decay mode**

  This is a five-digit integer number that codifies the different B_{tag} decays. It is shown in Figure 4.13.

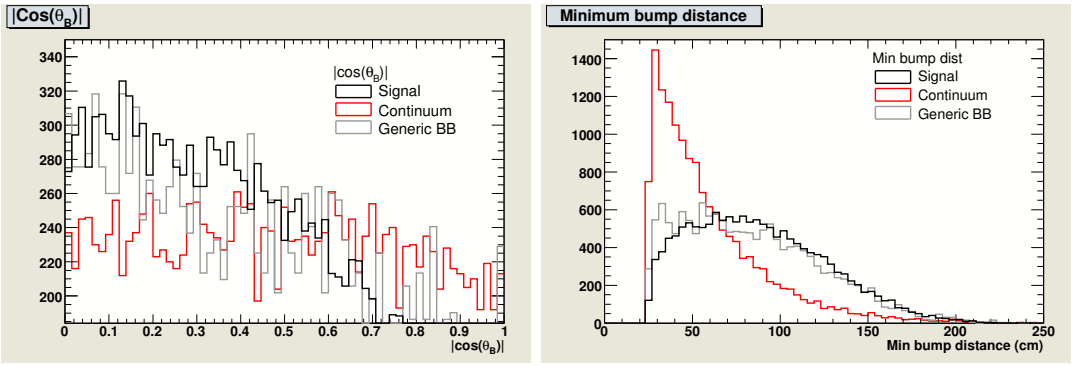

Figure 4.12. $|\cos(\theta_B)|$ (left) and minimum bump distance on the EMC surface of the high energy photon from any other cluster (right) for signal, continuum and generic $B\bar{B}$ events.

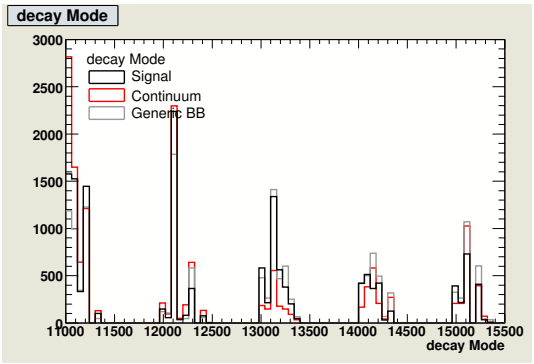

Figure 4.13. Decay mode of the B_{tag} (bottom) for signal, continuum and generic $B\bar{B}$ events.

- **CLEO energy flow cones**

  The CLEO energy cones are defined as the sum of the energies of reconstructed particles with momenta within defined angular regions around the high energy photon. We define two sets of cones:

  – all particles in the event, with energies calculated in the $\Upsilon(4S)$ frame;

  – decay products of $B_{\rm sig}$, with energies calculated in the $B_{\rm sig}$ rest frame.

  Each set is composed of eighteen cones, each one of them defined in a ten-degree angular region around the high energy photon. The photon is excluded from the sum.

  Since $q\overline{q}$ events are distributed back-to-back in the $\Upsilon(4S)$ frame, we expect most of the their energy to be in those cones defined in angular regions closer – or in the opposite side of – the photon direction. On the other hand, B events are isotropically distributed, henceforth the cones should be approximately equally populated.

  The situation is different for the cones composed of $B_{\rm sig}$ particles in the $B_{\rm sig}$ frame. For $B \to X_s\gamma$ decays, the photon and the X_s are distributed back-to-back in the $B_{\rm sig}$ frame. For signal events, we expect the cones in the opposite side of the photon direction to be more energetic and the cones near the photon direction to have little energy. All the distributions are shown in Figures 4.14 to 4.17.

4.4.2 Training of a Boosted Decision Tree

We use forty-four variables to train the MVA discriminant: thirty-six are CLEO energy cones and the rest are related to spatial distributions of the particles. Since the energy cones present higher correlations among them with respect to the other variables, the construction of the MVA variables is divided in two consecutive steps:

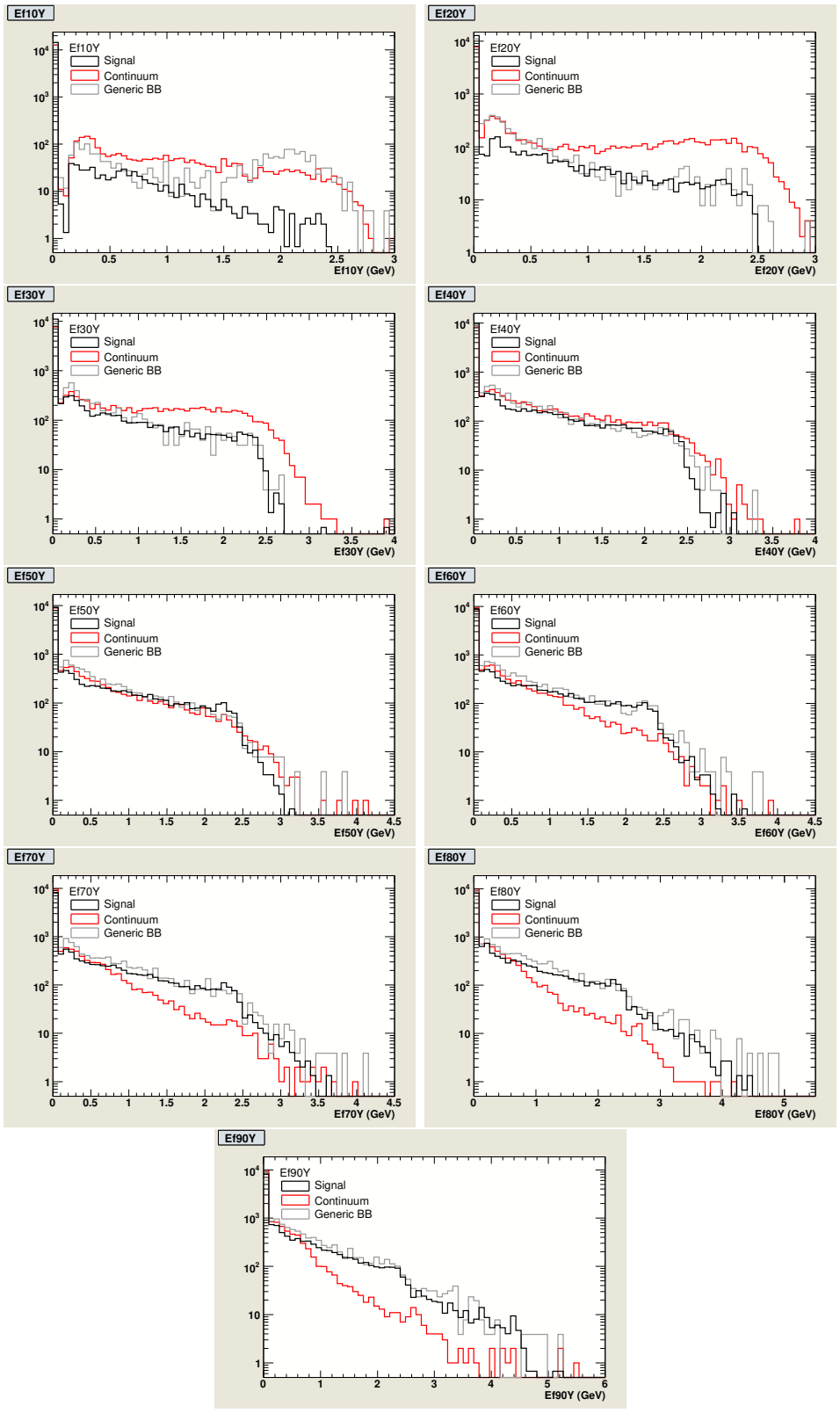

Figure 4.14. CLEO energy cones calculated with all the particles with energies in the $\Upsilon(4S)$ frame. They lie between 0 and 90 degrees from the photon direction. The distributions are drawn in logarithmic scale because of the high peak at zero.

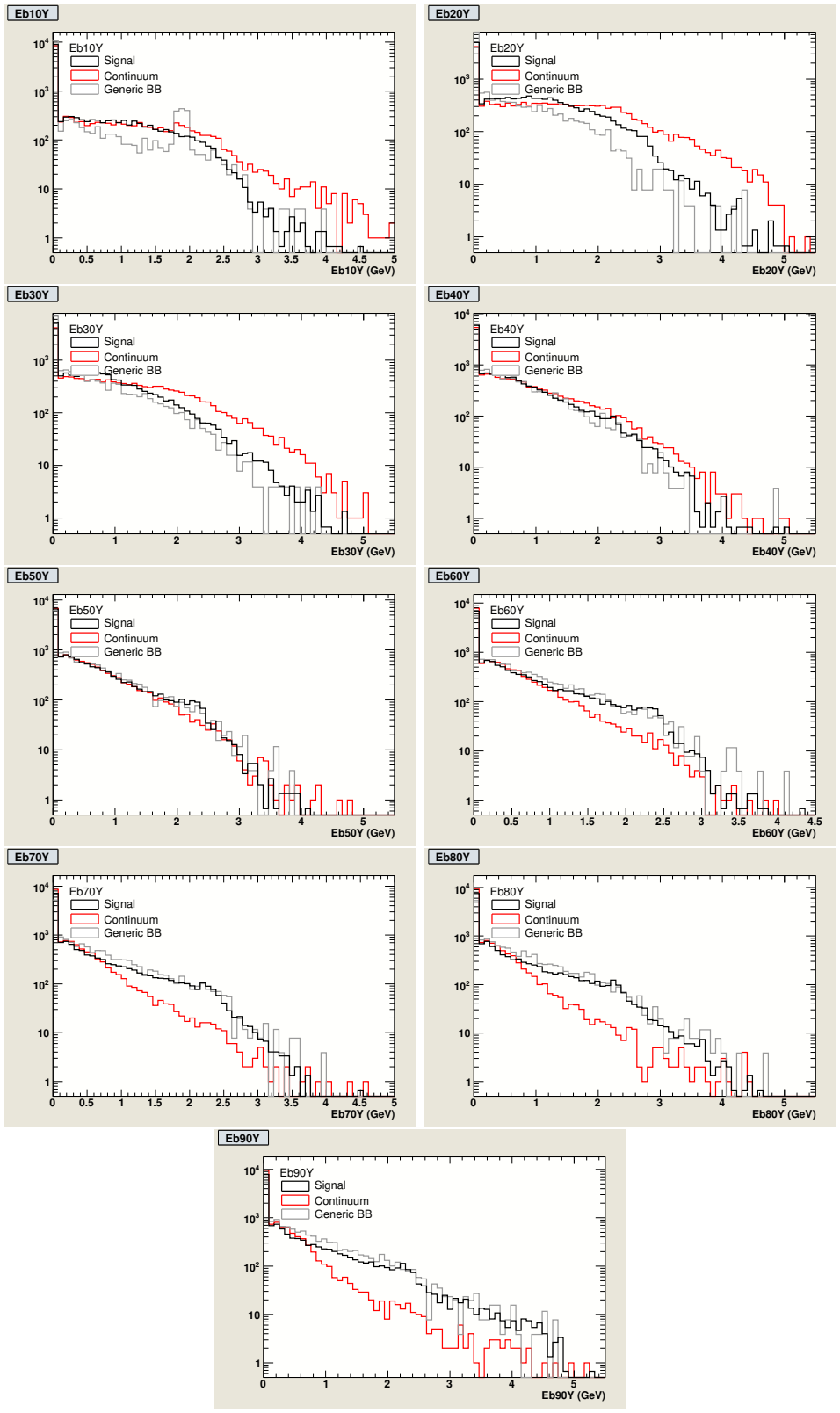

Figure 4.15. CLEO energy cones calculated with all the particles with energies in the $\Upsilon(4S)$ frame. They lie between 0 and 90 degrees from the direction opposite to the photon. The distributions are drawn in logarithmic scale because of the high peak at zero.

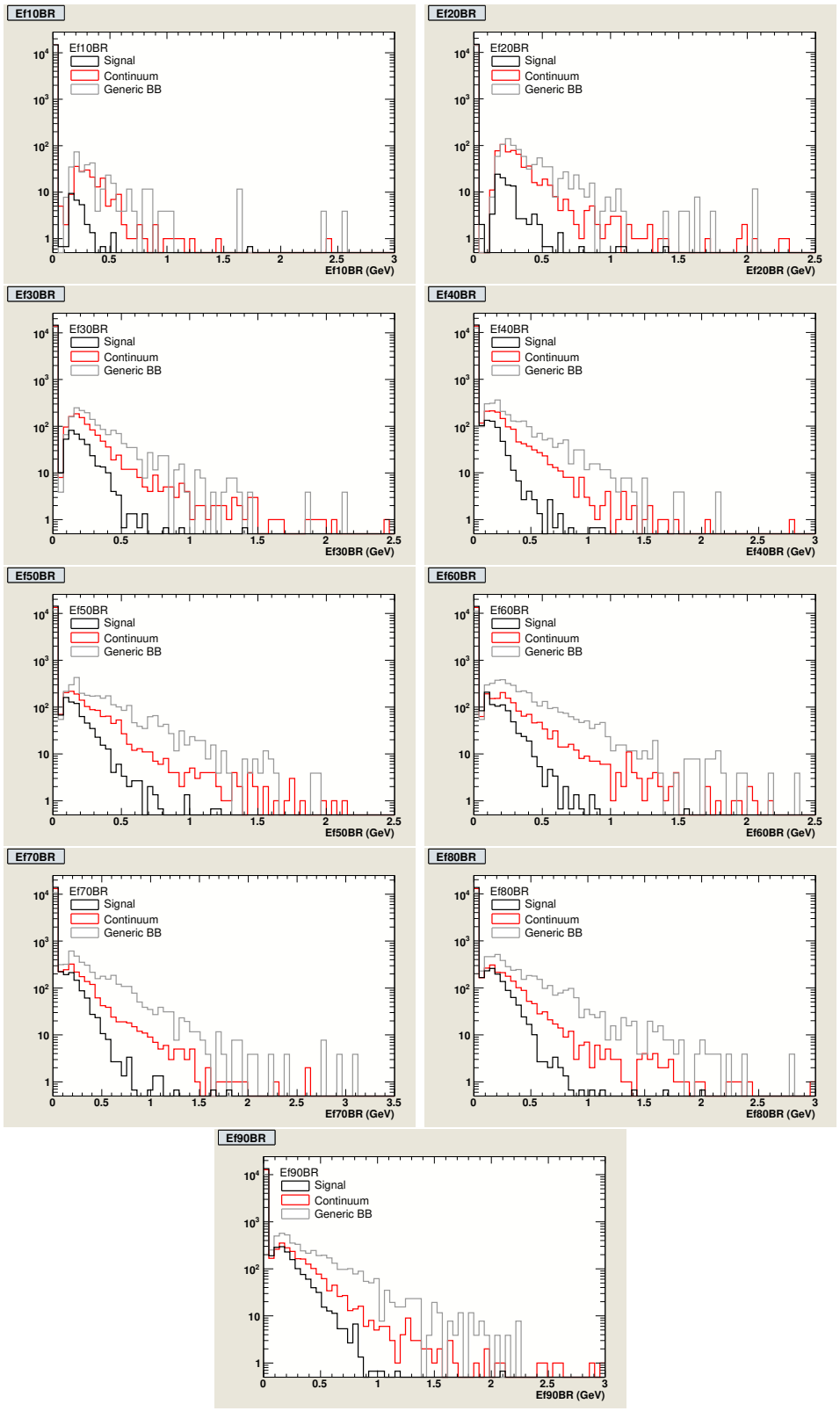

Figure 4.16. CLEO energy cones calculated with decay products of $B_{\rm sig}$, with energies in the $B_{\rm sig}$ frame. They lie between 0 and 90 degrees from the photon direction. The distributions are drawn in logarithmic scale because of the high peak at zero.

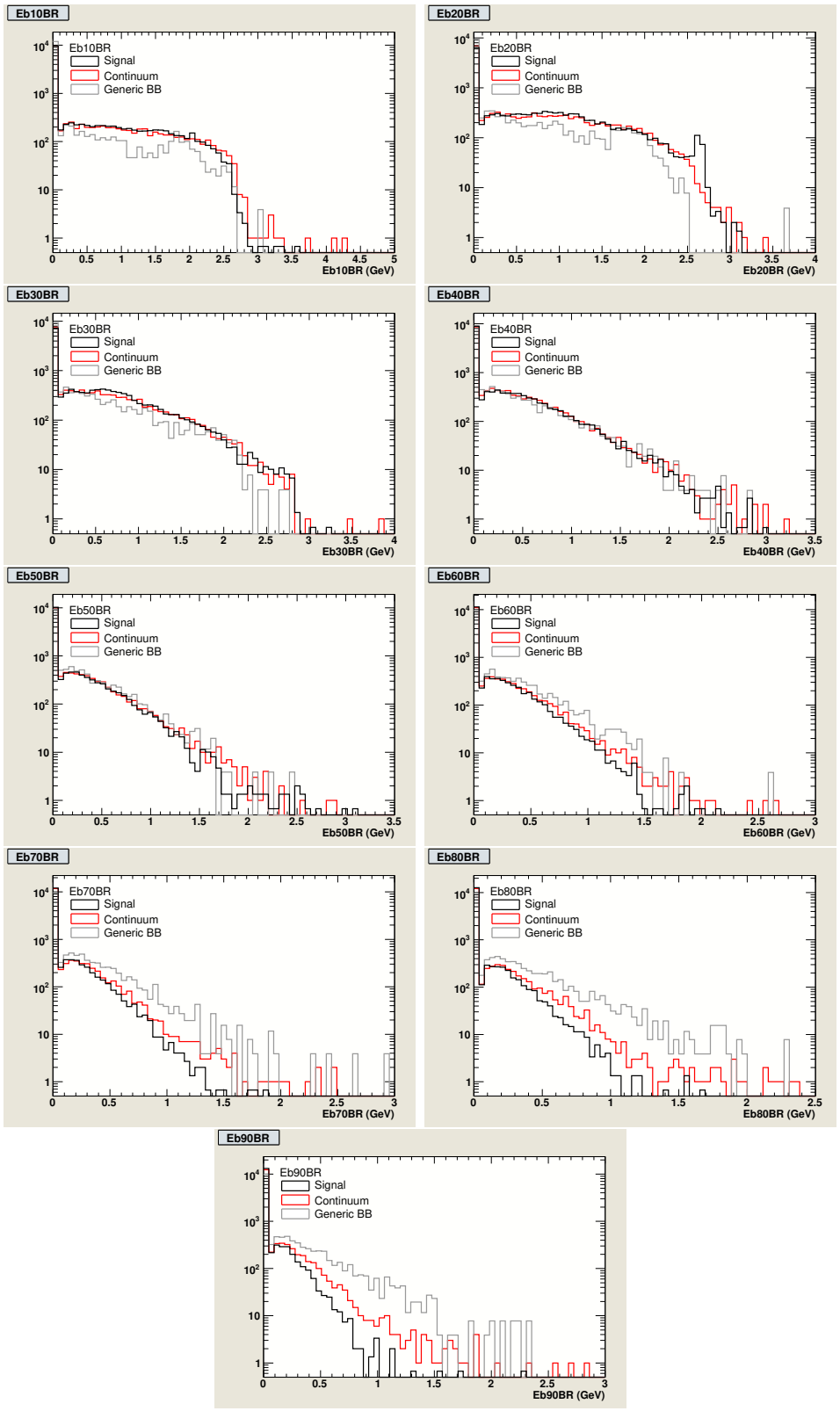

Figure 4.17. CLEO energy cones calculated with decay products of $B_{\rm sig}$, with energies in the $B_{\rm sig}$ frame. They lie between 0 and 90 degrees from the direction opposite to the photon. The distributions are drawn in logarithmic scale because of the high peak at zero.

- the first step is to build five different MVA variables (either a Fisher or a likelihood) by using only the energy cones;

- the second step is to combine the five MVA variables with all the other input variables into the final MVA discriminant.

We split the total available MC samples in three blocks. We use $\sim$ 7500 signal and continuum MC events for the training and the same numbers for the validation of the MVA variables. This choice represents $\sim$ 70 % of the available continuum MC sample. The remaining events are used to construct the MVA variable used in the fit, as described in the next section. The samples for the training and for the validation are statistically independent of each other, in order to avoid a potential bias in the MVA discriminant. The samples are constituted only of events which satisfy the requirement on the photon energy, $E_\gamma > 2.1$ GeV and those requirements on $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ events, listed in Table 4.12.

The first step of the training consists of combining the CLEO energy cones in several MVA discriminants, in order to enhance their discriminating power. In detail, they are:

- a Fisher for cones defined in the $\Upsilon(4S)$ frame, with all particles in the event, oriented along the direction of the photon;

- a Fisher for cones defined in the $\Upsilon(4S)$ frame, with all particles in the event, oriented in the opposite direction of the photon;

- a likelihood for cones defined in the $B_{\rm sig}$ frame, with particles coming from $B_{\rm sig}$ decays, oriented along the direction of the photon;

- a Fisher for cones defined in the $B_{\rm sig}$ frame, with particles coming from $B_{\rm sig}$ decays, oriented in the opposite direction of the photon;

- a Fisher built using all thirty-six cones.

The second step is the training of the actual variable that is used in the maximum likelihood fit. Thirteen variables (five associated to the CLEO energy cones and eight related to the spatial distribution of the event) compose the final MVA discriminant. All the input variables, shown in Figure 4.18, are scaled into a $[0 \div 1]$ range. As already mentioned, we train three different variables: a Fisher, a neural network (MLP, i.e. MultiLayer Perceptron) and a boosted decision tree (BDT). The three distributions are shown in Figure 4.19. The correlations matrix for signal events and for background events are shown in Figures 4.20 and 4.21.

Even if MLP has a slightly better discriminating power, as seen in the background rejection versus signal efficiency curve, Figure 4.22, we choose to use BDT in the maximum likelihood fit because it is easier to find a functional form to its shape, for an easier fit parameterization (see Section 4.5).

4.4.3 Construction of a Boosted Decision Tree on MC samples

The final step in the construction of the BDT is to apply the weights obtained during the training phase to a statistically independent set of MC events. As already mentioned, ~ 70 % of $q\overline{q}$ MC events are used to train the BDT, therefore continuum is the smallest available MC sample. We report the total number of events for $E_\gamma > 1.9$ GeV, for $E_\gamma > 2.1$ GeV and the corresponding equivalent integrated luminosity for each MC sample in Table 4.13. Figure 4.23 shows the BDT distribution for the

MC sample	Lumi (fb^{-1})	# Events $E_\gamma > 1.9$ GeV	# Events $E_\gamma > 2.1$ GeV
signal	2.0×10^5	99470	70653
continuum	449.0	16127	6340
generic $B\bar{B}$	1.4×10^3	13411	3858

Table 4.13. Total number and corresponding equivalent integrated luminosity for signal, continuum and generic $B\bar{B}$ MC events, with $E_\gamma > 1.9$ GeV and $E_\gamma > 2.1$ GeV, remaining after the training of and used for the construction of the BDT.

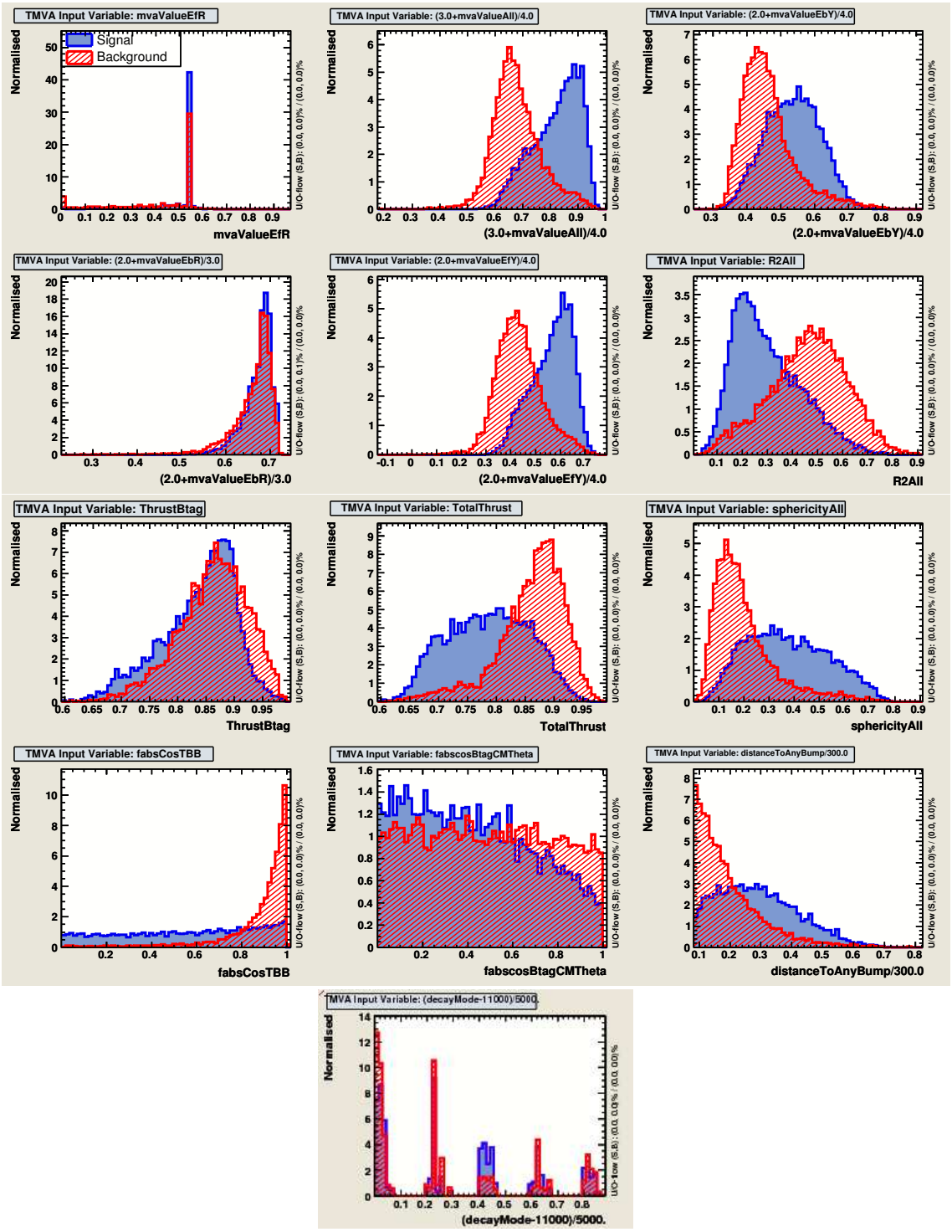

Figure 4.18. Distribution of the thirteen input variables used to build the final *BDT*. They include the five MVA variables built in the first step (the first five figures) and all the other variables previously shown. All distributions are scaled into a $[0 \div 1]$ range.

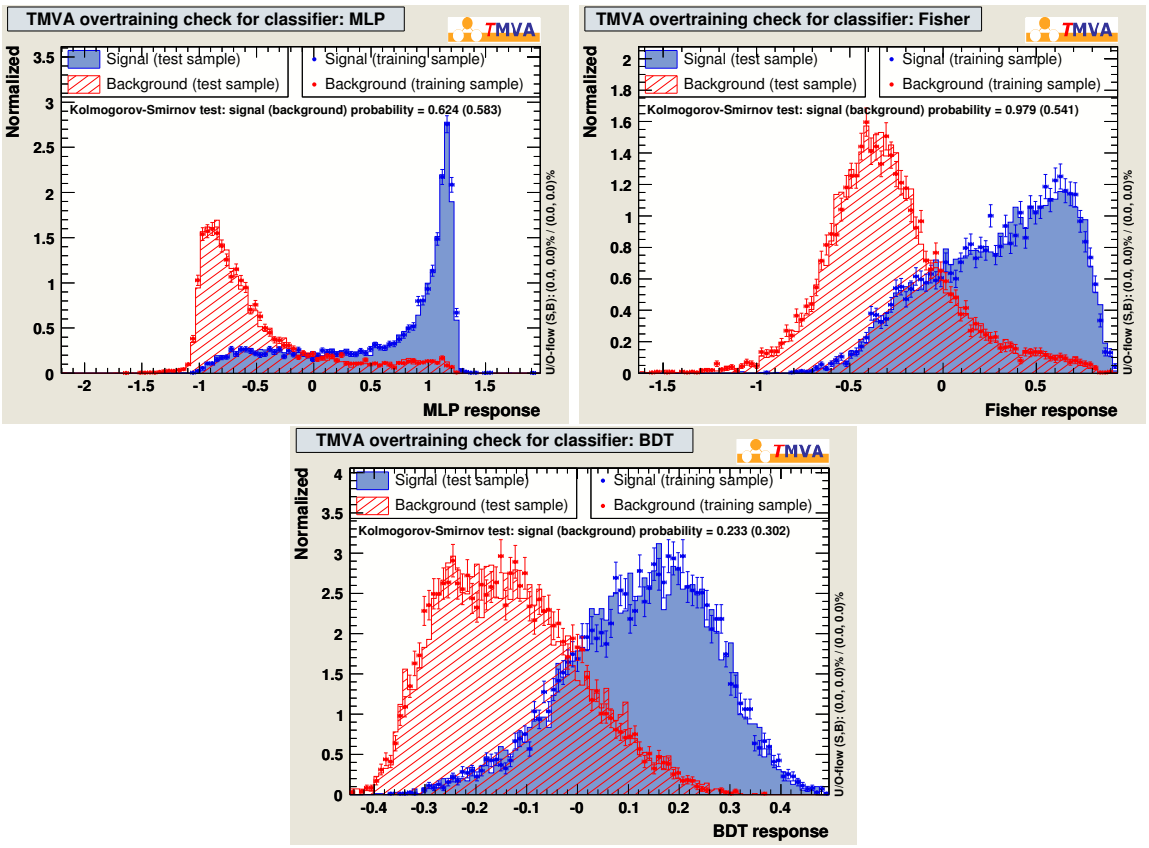

Figure 4.19. Distribution of MLP (top-left), Fisher (top-right) and *BDT* (bottom) for signal (blue) and background (red). The dots are the training sample of each MVA variable, the continuous curves are the corresponding validations samples. Also shown is the probability of the Kolmogorov-Smirnov test.

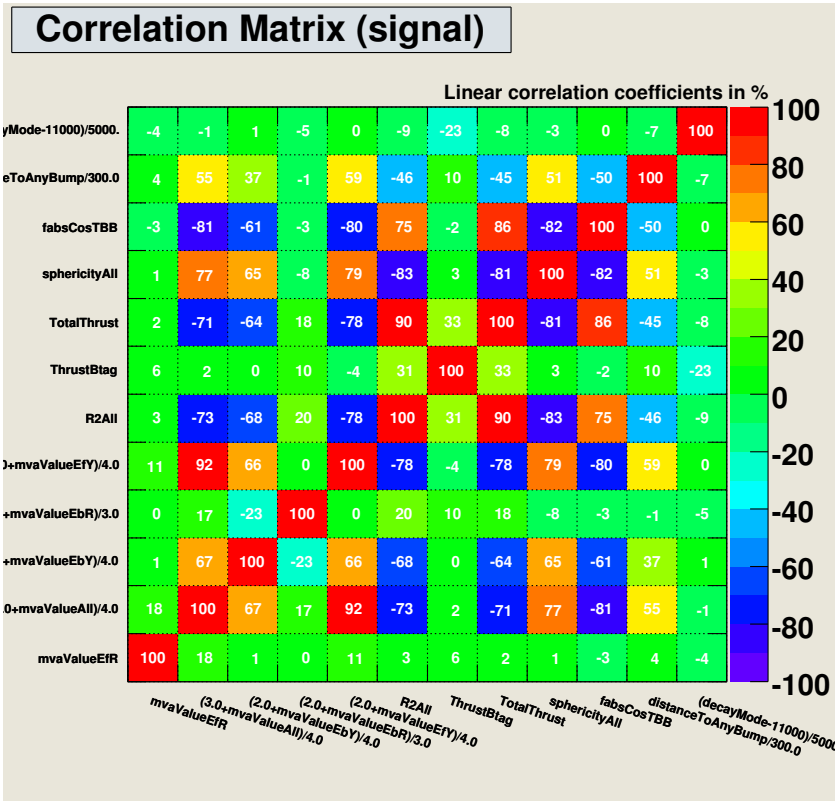

Figure 4.20. Correlation matrix for the thirteen normalized input variables used to build the *BDT*, for signal events only.

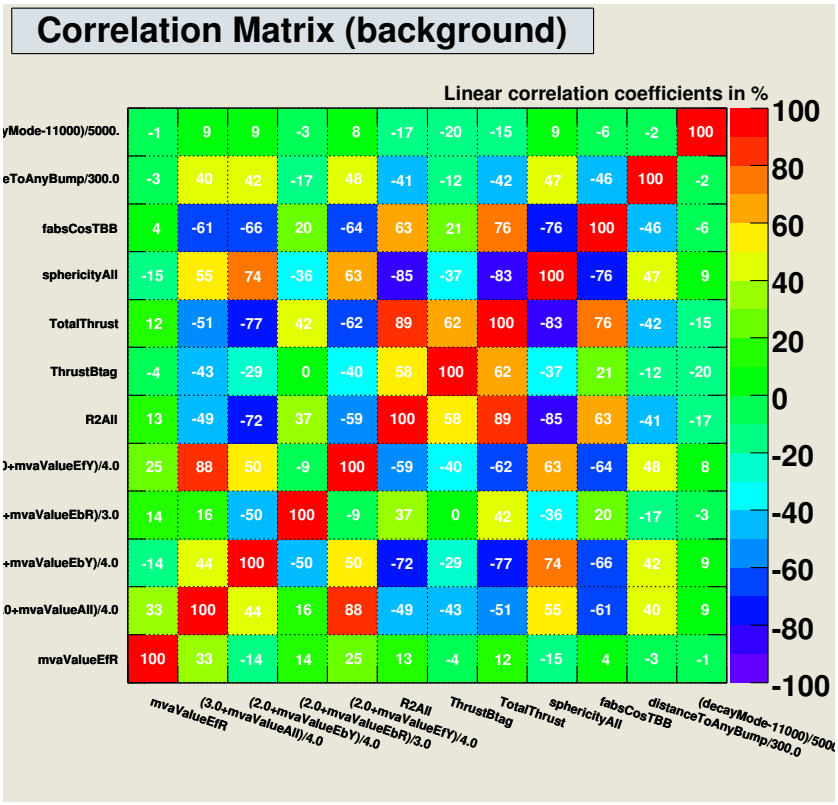

Figure 4.21. Correlation matrix for the thirteen normalized input variables used to build the *BDT*, for background events only.

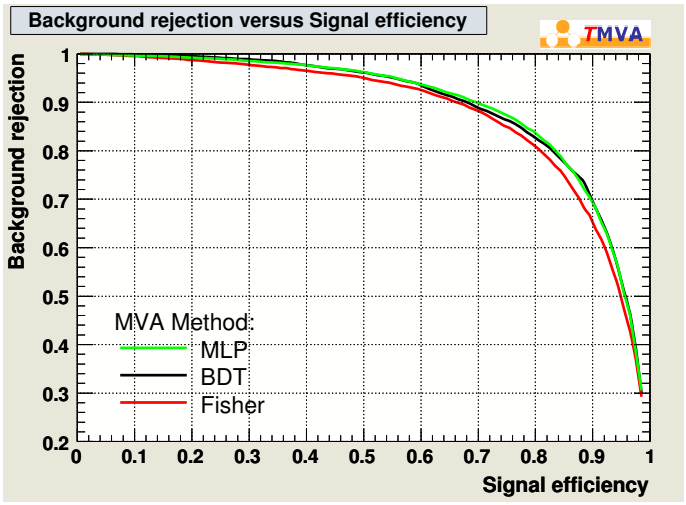

Figure 4.22. Background rejection versus signal efficiency curve for MLP, *BDT* and Fisher discriminant.

three MC samples. As already mentioned, these distributions are obtained on events with a photon cut of $E_\gamma > 2.1$ GeV and with requirements aimed at eliminating $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ events.

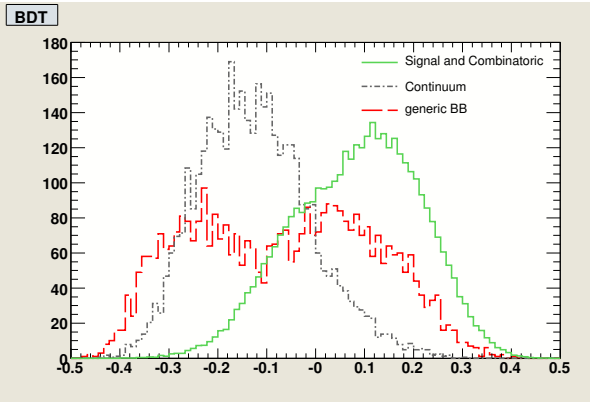

Figure 4.23. *BDT* distributions for signal an combinatoric events (green solid curve), continuum events (grey dot-dashed curve) and peaking and combinatoric components of generic $B\bar{B}$ (red dashed curve).

The off-resonance sample is useful to check the validity of the *BDT* distribution obtained on the continuum sample. Figure 4.24 is the comparison of the *BDT* for continuum and off-resonance data.

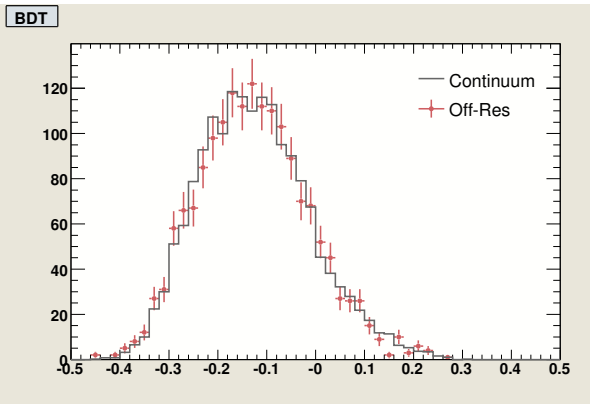

Figure 4.24. *BDT* distribution for off-resonance data (dots) and continuum MC events.

4.5 Maximum likelihood fit

We extract the *CP* asymmetry with an unbinned, multidimensional maximum likelihood fit. It is constructed with two variables: $m_{\rm ES}$, calculated from $B_{\rm tag}$ decay products, and BDT, described in the previous section. The corresponding probability density functions are estimated from the MC simulations, as explained in section 4.5.1. There is one signal and four background components: combinatoric, continuum, peaking (in $m_{\rm ES}$) and combinatoric generic $B\bar{B}$. We fit events with $E_\gamma > 2.2\,{\rm GeV}$. The choice of the best cut is based on toy MC studies, as explained in Section 4.5.4. We extract A_{CP} by including the flavor of $B_{\rm tag}$ (an integer number corresponding to either a B or a $\bar{B}$) as a fit category in the likelihood. The full expression is:

$$
\begin{aligned}
\mathcal{L} \quad = \quad & \frac{e^{-(N_{\rm sig}+N_{\rm comb}+N_{q\overline{q}}+N_{B\bar{B}})/N}}{N\sqrt{(N_{\rm sig}+N_{\rm comb}+N_{q\overline{q}}+N_{B\bar{B}})!}} \cdot \prod_{i=1}^{N} \Big\{ \\
N_{\rm sig}\cdot \quad & \Big[\frac{1-A_{\rm sig}}{2}+\frac{1+A_{\rm sig}}{2}\Big]\cdot P_{\rm sig}(m_{\rm ES})_i \cdot P_{\rm sig}(BDT)_i + \\
N_{\rm comb}\cdot \quad & \Big[\frac{1-A_{\rm comb}}{2}+\frac{1+A_{\rm comb}}{2}\Big]\cdot P_{\rm comb}(m_{\rm ES})_i \cdot P_{\rm comb}(BDT)_i + \\
N_{q\overline{q}}\cdot \quad & \Big[\frac{1-A_{q\overline{q}}}{2}+\frac{1+A_{q\overline{q}}}{2}\Big]\cdot P_{q\overline{q}}(m_{\rm ES})_i \cdot P_{q\overline{q}}(BDT)_i + \\
N_{B\bar{B}p}\cdot \quad & \Big[\frac{1-A_{B\bar{B}p}}{2}+\frac{1+A_{B\bar{B}p}}{2}\Big]\cdot P_{B\bar{B}p}(m_{\rm ES})_i \cdot P_{B\bar{B}p}(BDT)_i + \\
N_{B\bar{B}c}\cdot \quad & \Big[\frac{1-A_{B\bar{B}c}}{2}+\frac{1+A_{B\bar{B}c}}{2}\Big]\cdot P_{B\bar{B}c}(m_{\rm ES})_i \cdot P_{B\bar{B}c}(BDT)_i \Big\}, \qquad (4.3)
\end{aligned}
$$

where:

- N is the total number of events observed in the fitted sample;
- $N_{\rm sig}$ and $A_{\rm sig}$ are the yield and A_{CP} for signal events;
- $N_{\rm comb}$, $N_{q\overline{q}}$, $N_{B\bar{B}p}$, $N_{B\bar{B}c}$, $A_{\rm comb}$, $A_{q\overline{q}}$, $A_{B\bar{B}p}$ and $A_{B\bar{B}c}$ are yields and *CP* asymmetries for background events;
- $P(X)_i$ is the PDF of the observable X for the event i.

We use the same PDF for events tagged as B and events tagged as $\bar{B}$.

The likelihood is built in order to minimize the error on the *CP* asymmetry for signal decays, since it is the goal of this analysis. $m_{\rm ES}$ has a big discriminating power for signal events against combinatoric and continuum events, but it is not able to distinguish between these two background components. *BDT* introduces additional discriminating power against continuum events, by construction. However, we do not have a real discriminating power against other B decays. This introduces a correlation between signal and generic $B\bar{B}$ (in both yields and *CP* asymmetries), therefore we fix the latter component to its MC expected values in the fit. We fit for yield and *CP* asymmetry of signal, combinatoric and continuum events. We also let the fit determine the parameterization of $q\overline{q}$ PDF's. Special attention is dedicated to the *BDT* of the $B\bar{B}$ components. We fix all the other shape parameters to the values estimated from MC.

4.5.1 Parameterization of the fit PDF's

We parameterize the PDF's by using unbinned ML fits to MC samples of all the fit components. We use only events with $E_\gamma > 2.2$ GeV. The values of all the fit parameters obtained in the next section are listed in Table 4.14.

4.5.1.1 $m_{\rm ES}$

Signal and generic $B\bar{B}$ events peak near the B mass and are described using a Crystall Ball function:

$$f(x;\alpha,\beta,\mu,\sigma) = \begin{cases} e^{-(\mu-\sigma)^2/2\sigma^2} & \text{if } (x-\mu)/\sigma > -\alpha, \\ \left(\frac{\beta}{|\alpha|} - |\alpha| - \frac{x-\mu}{\sigma}\right)^{-\beta} \left(\frac{\beta}{|\alpha|}\right)^{\beta} e^{-\alpha^2/2} & \text{otherwise} \end{cases} \tag{4.4}$$

where μ is the mean value of the distribution, σ is the RMS of the Gaussian component, α and β describe the non-Gaussian left tail. All the $m_{\rm ES}$ parameters for signal

and generic $B\bar{B}$ components are fixed in the fit. The combinatoric and continuum $m_{\rm ES}$ components are parameterized with an ARGUS [60] function: We fix the combinatoric shape parameters, but we float the corresponding $q\overline{q}$. In order to perform tests on the fit, such as toy MC studies, we use the $q\overline{q}$ parameters obtained from the simulation as a reference. The shapes are shown in Figure 4.25.

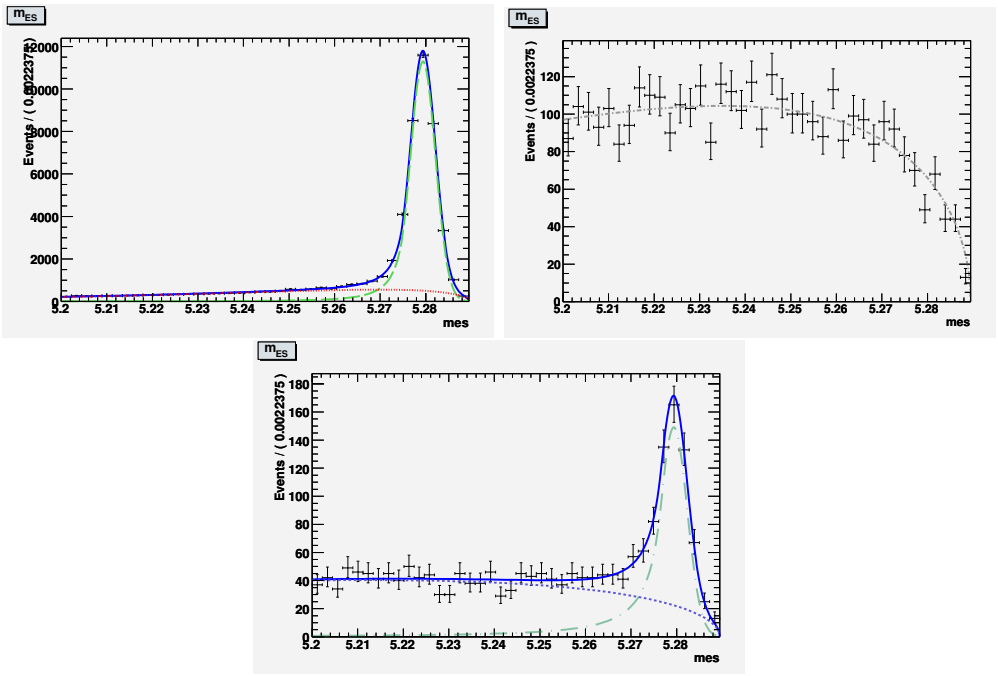

Figure 4.25. Fit to the $m_{\rm ES}$ distribution for signal and combinatoric (left), $q\overline{q}$ (right) and the two components of generic $B\bar{B}$ (bottom) obtained from MC samples.

4.5.1.2 *BDT*

All the *BDT*'s used in this analysis are parameterized with a sum of two Gaussians:

$$f(x) = f \cdot e^{-(x-\mu_1)^2/2\sigma_1^2} + (1-f) \cdot e^{-(x-\mu_2)^2/2\sigma_2^2} \tag{4.5}$$

Thus each *BDT* is characterized by five parameters.

While we possess a MC sample of pure continuum events, the situation is slightly more complicated for the other four components. Combinatoric events are mainly

mis-reconstructed photons in the hadronic-tag side of the event, therefore the MC sample that we use for signal inevitably contains combinatoric events as well. In order to effectively parameterize the signal and combinatoric BDT shapes, we use the discriminating power of $m_{\rm ES}$. The BDT parameterization is performed in two steps.

- We determine the $m_{\rm ES}$ shapes for the signal and combinatoric PDF's, as explained in the previous section;
- we fit the signal MC sample for $m_{\rm ES}$ and BDT simultaneously by keeping all the $m_{\rm ES}$ parameters and the two yields fixed to the values obtained in the first step.

We apply the same procedure for the two generic $B\bar{B}$ components.

Signal, combinatoric and continuum parameters are fixed, to the values obtained from the simulation, in the nominal fit. The only exception is the ARGUS shape parameter of continuum, which is left free. Since it is not easy to validate the shapes for generic $B\bar{B}$ events with a data sample, we decide to float their parameters. However, given a high correlation between the signal and the peaking $B\bar{B}$ components, we float the peaking BDT parameters only in the limited range of one standard deviation, whereas we leave the non-peaking BDT completely free. As for the peaking BDT parameters, we multiply the BDT PDF by five Gaussians, each one of them built with mean and RMS equal to the value and the error of the corresponding BDT parameter – estimated with MC – respectively.

Figure 4.26 shows the shapes for all the components.

4.5.2 Correlation between the fit variables

We study the correlation between BDT and $m_{\rm ES}$ for all the components in the fit. We use the available MC sample for each component. Profile plots are used, shown in

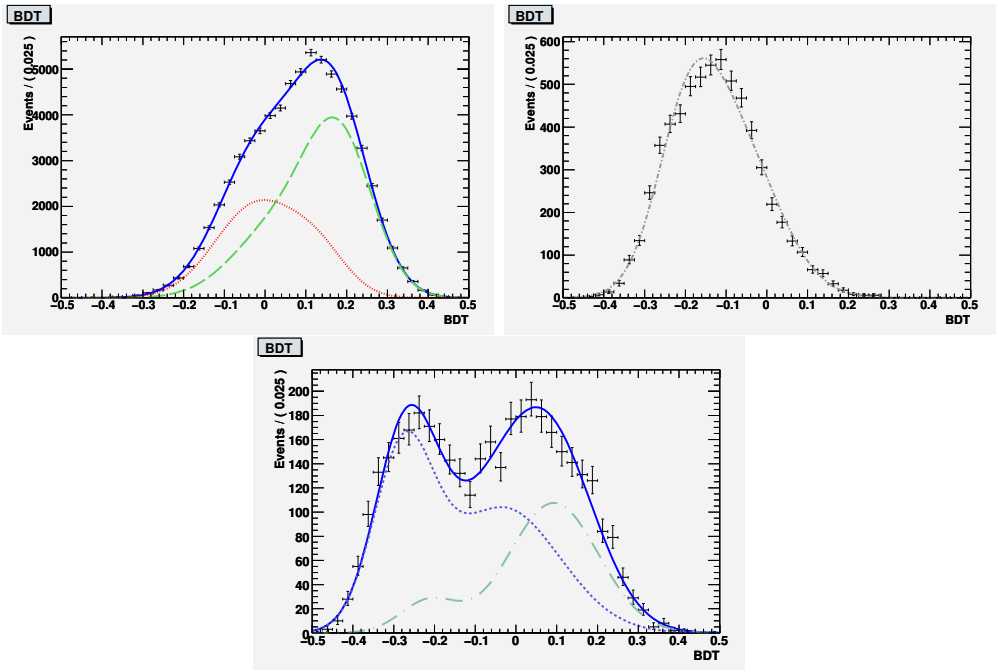

Figure 4.26. The top-left figure is the fit to the BDT distribution for signal (right dashed curve) and combinatoric (left curve); the top-right figure is for $q\overline{q}$ and the bottom figure is for the two components of generic $B\bar{B}$ (the peaking is the right hand dot-dashed curve) obtained from MC samples.

Parameter	Value
$N_{B\bar{B}p}$	210
$N_{B\bar{B}c}$	438
SIG $m_{\rm ES}$ μ	5.279400 ± 0.000021
SIG $m_{\rm ES}$ α	1.415 ± 0.051
SIG $m_{\rm ES}$ σ	0.002744 ± 0.000025
SIG $m_{\rm ES}$ N	2.55 ± 0.46
COMB $m_{\rm ES}$ ξ	-65.9 ± 2.2
$B\bar{B}_P$ $m_{\rm ES}$ μ	5.27940 ± 0.00021
$B\bar{B}_P$ $m_{\rm ES}$ σ	0.00294 ± 0.00024
$B\bar{B}_P$ $m_{\rm ES}$ α	1.00 ± 0.19
$B\bar{B}_P$ $m_{\rm ES}$ N	1.81 ± 1.51
$B\bar{B}_C$ $m_{\rm ES}$ ξ	-16.4 ± 8.0
SIG BDT μ_1	0.1851 ± 0.0025
SIG BDT σ_1	0.0826 ± 0.0010
SIG BDT μ_2	0.0267 ± 0.0092
SIG BDT σ_2	0.1061 ± 0.0033
SIG BDT f	0.6496 ± 0.030
COMB BDT μ_1	-0.0274 ± 0.0066
COMB BDT σ_1	0.1000 ± 0.0028
COMB BDT μ_2	0.1322 ± 0.0068
COMB BDT σ_2	0.0746 ± 0.0030
COMB BDT f	0.7135 ± 0.041
$q\overline{q}$ BDT μ_1	$-0.106960000 \pm 0.000000013$
$q\overline{q}$ BDT σ_1	$(7.0 \pm 6.0) \times 10^{-9}$
$q\overline{q}$ BDT μ_2	-0.1202 ± 0.0018
$q\overline{q}$ BDT σ_2	0.1111 ± 0.0013
$q\overline{q}$ BDT f	0.00023 ± 0.00025
$B\bar{B}_P$ BDT μ_1	-0.228 ± 0.012
$B\bar{B}_P$ BDT σ_1	0.0619 ± 0.0092
$B\bar{B}_P$ BDT μ_2	0.0829 ± 0.0099
$B\bar{B}_P$ BDT σ_2	0.1246 ± 0.0070
$B\bar{B}_P$ BDT f	0.209 ± 0.033

Table 4.14. Summary of the fit parameters, obtained with Monte Carlo samples of all the fit components. We list all and only fixed and peaking $B\bar{B}$ BDT parameters. The requirement of $E_\gamma > 2.2$ GeV is imposed.

Figures 4.27 and 4.28. The corresponding correlation factors are shown in Table 4.15.

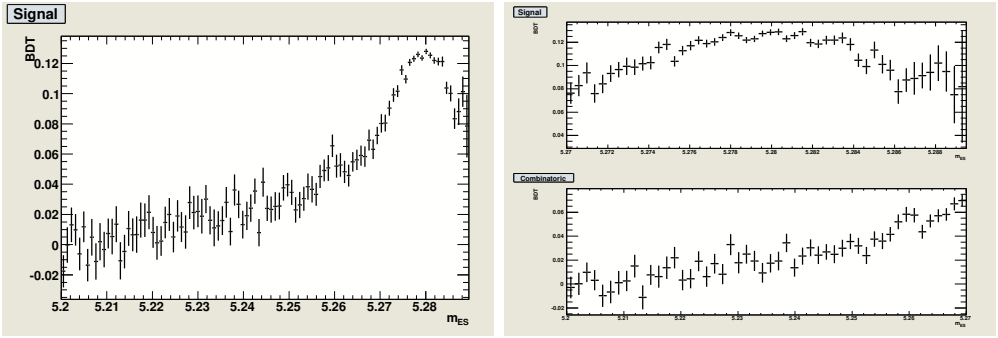

Figure 4.27. Profile plot of BDT vs $m_{\rm ES}$ distributions for signal and combinatoric events, taken from the corresponding MC sample. The requirement $E_\gamma > 2.2$ GeV is applied. On the left we plot over the whole $m_{\rm ES}$ fit range ($5.2 \div 5.2895$ GeV/c^2), whereas on the right we restrict to the peaking region ($m_{\rm ES} > 5.27$ GeV/c^2) for the top figure and to the sideband region ($m_{\rm ES} \leq 5.27$ GeV/c^2) for the bottom region.

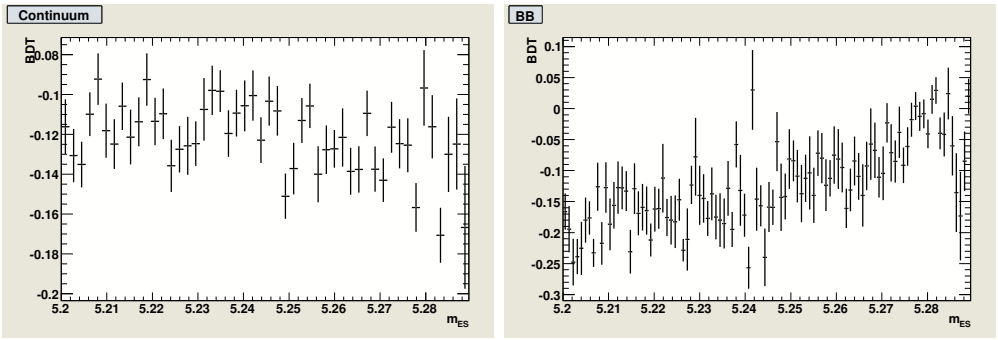

Figure 4.28. Profile plot of BDT vs $m_{\rm ES}$ distributions for continuum (left) and generic $B\bar{B}$ events (right), taken from the corresponding MC samples. The requirement $E_\gamma > 2.2$ GeV is applied.

4.5.3 Fit validation

We validate the fit using samples of MC events, coming from a full simulation or from toy experiments. In Table 4.16 we give the results of the fit on a sample of pure signal or background MC events.

Component	Correlation
Signal (full $m_{\rm ES}$ range)	0.29
Signal ($m_{\rm ES} > 5.27$ GeV/c^2)	0.028
Combinatoric ($m_{\rm ES} \leq 5.27$ GeV/c^2)	0.16
Continuum	-0.051
generic $B\bar{B}$	0.33

Table 4.15. Correlation factors between BDT and $m_{\rm ES}$ obtained on the profile plots shown in Figures 4.27 and 4.28 for the various fit components. The requirement of $E_\gamma > 2.2$ GeV is imposed.

Sample	$N_{\rm sig}$	A_{CP}(sig)
40576 Sig events 13349 Comb events	40603 ± 211	0.0009 ± 0.0052
2011 $B\bar{B}$ events	1.9 ± 1.2	20 ± 38
3682 $q\overline{q}$ events	0.0 ± 1.3	2 ± 27

Table 4.16. Fit result on $B\bar{B}$ and $q\bar{q}$ Monte Carlo samples.

In order to quantify any bias in the fit, we also perform a set of toy MC experiments in which we generate realistic samples (containing events coming from signal, continuum and generic B decays) according to the likelihood shape of each component and we fit them. All the generated samples correspond to a total equivalent integrated luminosity of 450 fb^{-1}. As in the nominal fit, we float $N_{\rm sig}$, $N_{\rm comb}$, $N_{\rm q\overline{q}}$, $A_{\rm sig}$, $A_{\rm comb}$, $A_{\rm q\overline{q}}$, the $q\overline{q}$ $m_{\rm ES}$ shape, the non-peaking $B\bar{B}$ BDT shape and we float the peaking $B\bar{B}$ BDT shape parameters within one standard deviation.

The pull and error distributions for the signal yield and CP asymmetry are shown in Figure 4.29. We estimate the biases by multiplying the means of the Gaussian fits to the pulls by the errors of the corresponding fits to the generated events, shown in Table 4.17.

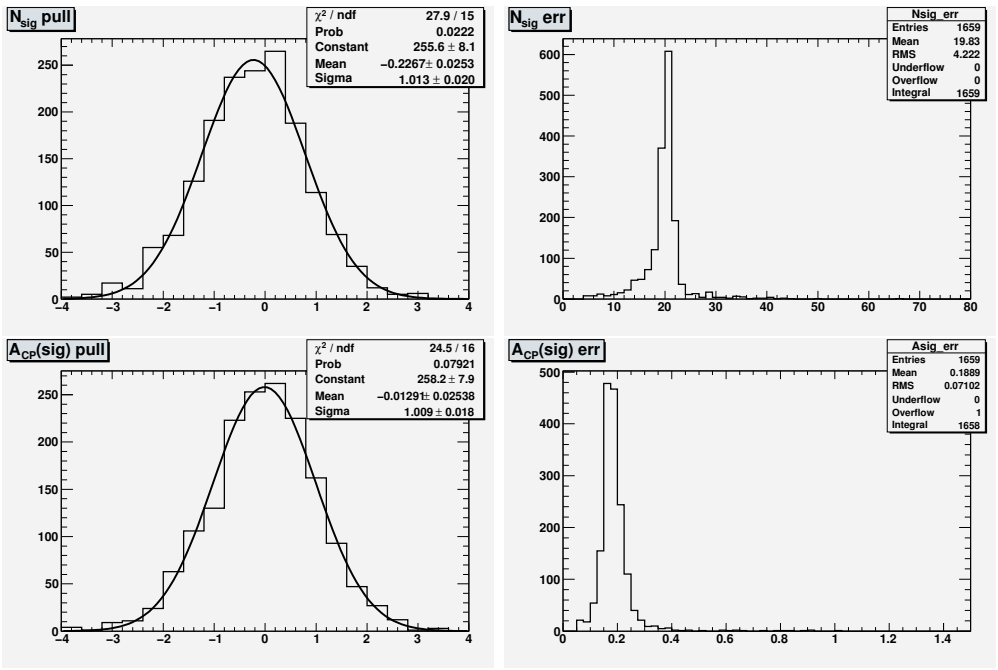

Figure 4.29. Pull (left) and error (right) distributions for signal yield and A_{CP} obtained from a set of 2000 toy MC esperiments performed with the expected inputs of Run1-6 dataset.

Parameter	Generated value	Bias
N_{sig}	112	-4.5
A_{CP}	0	-0.0024

Table 4.17. Results from 2000 toy MC experiments. We report the generated signal yield and A_{CP} and the corresponding bias obtained by multiplying the mean of the pull distribution by the error obtained from the fits. The generated sample corresponds to an expected total intergrated luminosity of 450 fb^{-1}.

$E_\gamma >$ (GeV)	N_{gen}	$N \pm \sigma(N)$	$A_{CP} \pm \sigma(A_{CP})$
1.9	318	315 ± 67	0.00 ± 0.16
2.0	296	318 ± 49	0.00 ± 0.12
2.1	261	278 ± 35	0.01 ± 0.11
2.2	224	219 ± 25	0.00 ± 0.11
2.3	174	162 ± 20	0.00 ± 0.12

Table 4.18. Toy MC experiments for different photon energy cuts. For each cut we report the number of generated signal events, the corresponding fitted number of events and the A_{CP}.

4.5.4 Choice of best photon energy cut

We choose the best cut on the minimum photon energy in order to minimize the experimental error on the signal yield and A_{CP}.

Since the variables used to train the BDT are not highly correlated with the photon energy, we expect that the PDF shapes do not change much by applying a different cut on the photon energy. However, a different cut on the photon energy changes the relative normalization of the peaking and combinatoric components for $B \to X_s\gamma$ and generic $B\bar{B}$ events. This is reasonable, since the probability of having a mis-reconstructed photon coming from hadronic $B_{\rm tag}$ decays is different if we change the minimum photon energy. This leads us to repeat the parameterization procedure, as explained in section 4.5.1, for the different cuts on E_γ. We scan the photon energy in steps of 100 MeV, from 1.9 GeV to 2.3 GeV. We report the PDF distributions for the different values of E_γ in Figures 4.30 to 4.32.

We perform toy MC experiments for each photon energy cut and report the values of yield and asymmetry for signal events in Table 4.18. Our choice is a cut of $E_\gamma >$ 2.2 GeV because, even if the error on A_{CP} is the same as the case of the looser cut at 2.1 GeV, there is an improvement in the error on the signal yield.

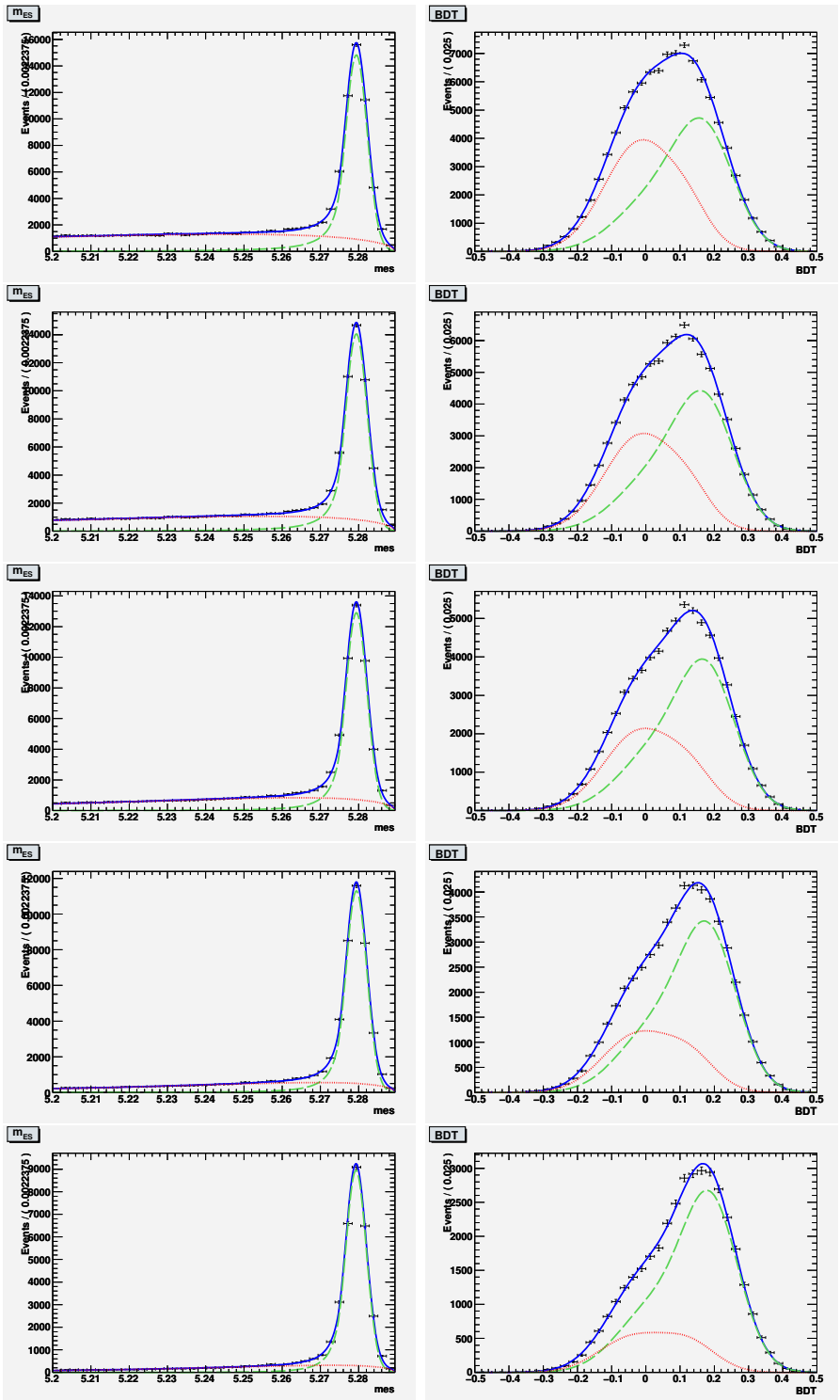

Figure 4.30. $m_{\rm ES}$ (left) and BDT (right) PDF's for signal (green dashed line) and combinatoric (red) components with $E_\gamma > 1.9$ (top row), $E_\gamma > 2.0$ (second row), $E_\gamma > 2.1$ GeV (third row), $E_\gamma > 2.2$ GeV (fourth row) and $E_\gamma > 2.3$ GeV (bottom row).

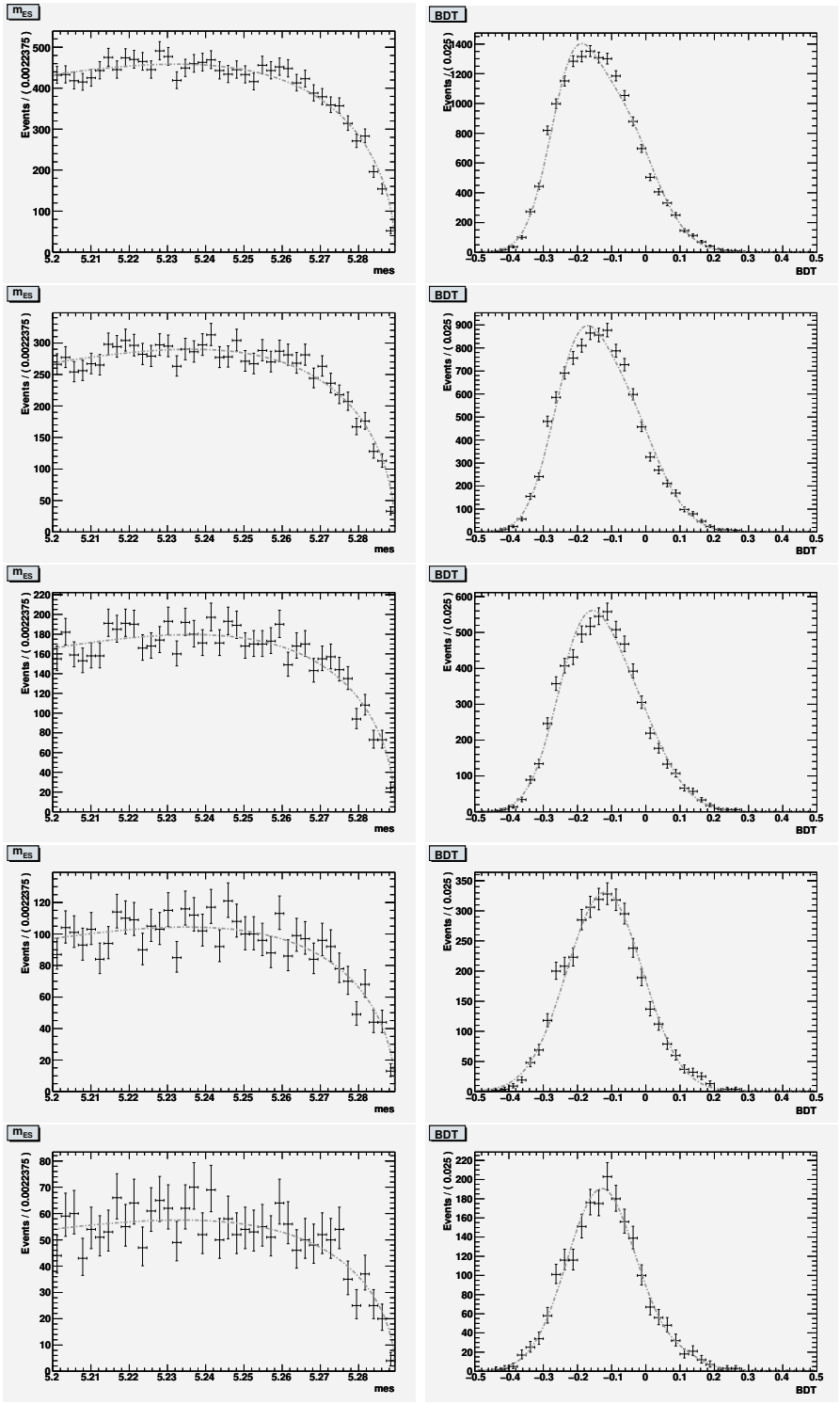

Figure 4.31. $m_{\rm ES}$ (left) and BDT (right) PDF's for the continuum component with $E_\gamma > 1.9$ (top row), $E_\gamma > 2.0$ (second row), $E_\gamma > 2.1$ GeV (third row), $E_\gamma > 2.2$ GeV (fourth row) and $E_\gamma > 2.3$ GeV (bottom row).

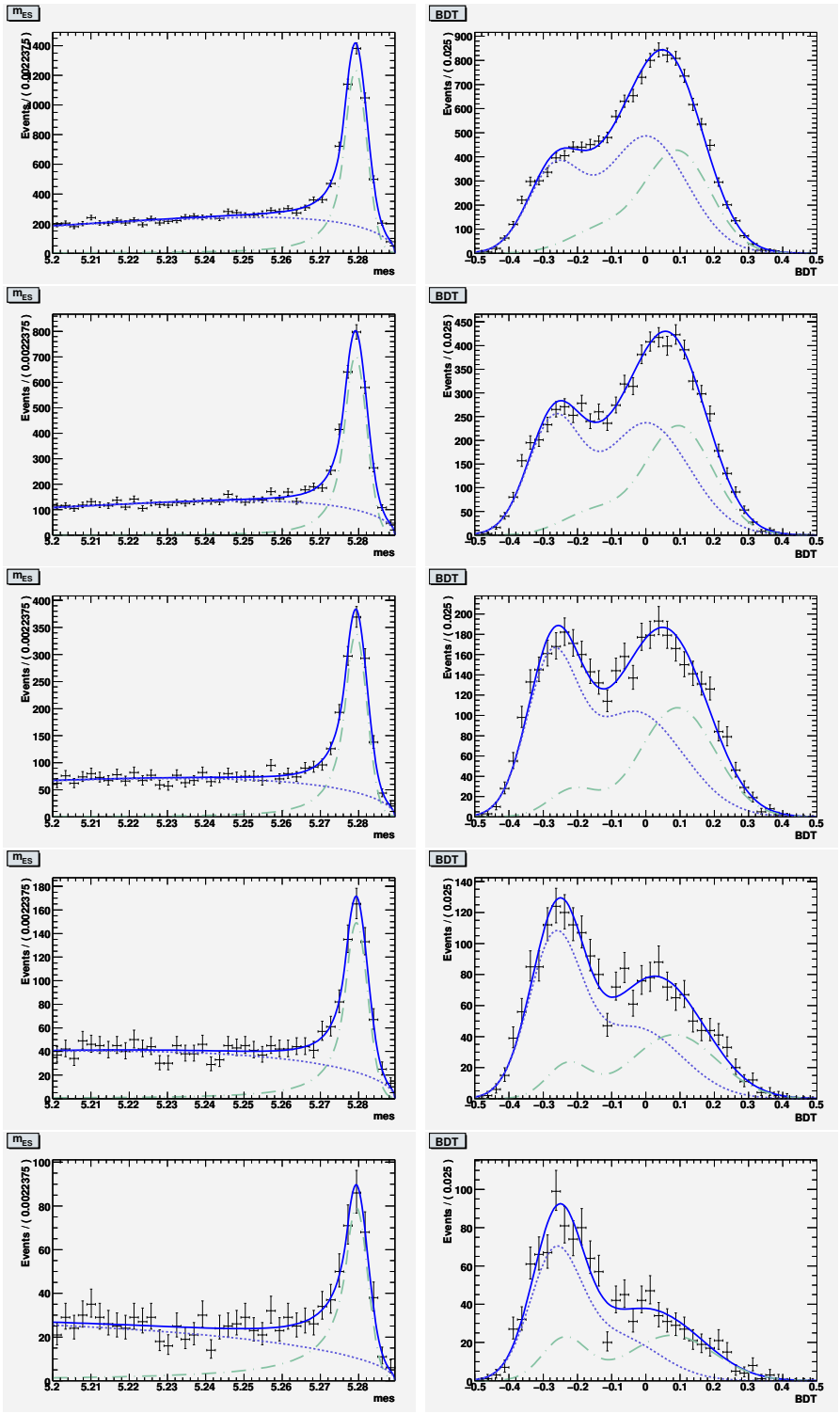

Figure 4.32. $m_{\rm ES}$ (left) and BDT (right) PDF's for peaking (right curve) and non-peaking (left curve) generic $B\bar{B}$ components with $E_\gamma > 1.9$ (top row), $E_\gamma > 2.0$ (second row), $E_\gamma > 2.1$ GeV (third row), $E_\gamma > 2.2$ GeV (fourth row) and $E_\gamma > 2.3$ GeV (bottom row).

$A_{CP}(BBp)$	$N \pm \sigma(N)$	$A_{CP} \pm \sigma(A_{CP})$
-0.5	109 ± 19	-0.41 ± 0.17
-0.4	108 ± 19	-0.33 ± 0.17
-0.3	108 ± 20	-0.25 ± 0.17
-0.2	107 ± 20	-0.17 ± 0.18
-0.1	107 ± 20	-0.08 ± 0.18
0.1	107 ± 20	0.08 ± 0.18
0.2	108 ± 19	0.17 ± 0.17
0.3	107 ± 20	0.25 ± 0.18
0.4	108 ± 20	0.33 ± 0.18
0.5	109 ± 20	0.42 ± 0.18

Table 4.19. Toy MC experiments with values of A_{CP} for the peaking generic $B\bar{B}$ component different than zero. We report the fitted signal yield and A_{CP}. for each value of the generated A_{CP} of the peaking $B\bar{B}$ component. A cut of $E_\gamma > 2.2$ GeV is assumed in this table.

4.5.5 Study of the bias due to fixing the $B\bar{B}$ component

As explained in Section 4.5, we fix the normalization and A_{CP} of the generic $B\bar{B}$ component in the fit, to the values obtained from the MC simulation and scaled to the Run1-6 integrated luminosity. We study the effect of this assumption on the signal yield and A_{CP}. We perform several sets of toy MC esperiments, with the nominal fit configuration, but with one exception. For each set of toys, we generate events in which the value of A_{CP} for the peaking $B\bar{B}$ component is different than zero in the generation PDF, whereas the corresponding value in the fit PDF is fixed to zero. We scan for different values of A_{CP} of the peaking $B\bar{B}$ component, ranging from -0.5 to 0.5, with step size of 0.1, and report the corresponding signal yield and A_{CP} in Table 4.19. The correlation between signal *CP* asymmetry and the $B\bar{B}$ one is clear from the table, which is a reason for fixing its value to zero in the nominal fit. The value of the *CP* asymmetry for $B\bar{B}$ events has been measured in the study of Reference [64], which presented the measurement of A_{CP} for the sum of several exclusive $b \to s\gamma$ decays. The systematic error on the signal A_{CP} due to $B\bar{B}$ is 0.011, for events with $E_\gamma > 2.2$ GeV.

Source	Syst. error
Signal shapes	0.011
Combinatoric shapes	0.0087
$q\overline{q}$ shapes	0.0068
$B\bar{B}$ shapes	0.023
Total	0.028

Table 4.20. Summary of systematic errors on the signal *CP* asymmetry due to the limited available MC statistics for the determination of signal and background PDF shapes parameters.

4.6 Systematics

The calculation of the *CP* asymmetry is affected by two main sources of systematic errors:

- systematic error on the yields;
- systematic error due to the $B\bar{B}$ background;
- mis-tag fraction due to $B^0 - \bar{B}^0 oscillations$;

In the following sections, we discuss each of these contributions.

4.6.1 Yields

The systematic error on the yields of B and $\bar{B}$ decays comes from the knowledge of the signal and background parameters. All the parameters are determined from MC simulations. The associated errors are used to evaluate the systematic ones, shifting each parameter by $\pm 1\sigma$ from its nominal value and quoting the difference on the *CP* asymmetry. A detailed list of the contributions is given in appendix B.

The summary of the systematic errors on the *CP* asymmetry is reported in Table 4.20.

4.6.2 Error due to $B\bar{B}$ background

The systematic error on the signal CP asymmetry due to $B\bar{B}$ background is reported in Section 4.5.5. It is equal to 0.011, for events with $E_\gamma > 2.2$ GeV.

4.6.3 Mis-tag fraction due to $B^0 - \bar{B}^0$ mixing

As already discussed in Section 1.6, the measurement of CP asymmetry for neutral B decays must be diluted by an irreducible source of mis-tagging, due to $B^0 - \bar{B}^0$ oscillations, as already shown in Eq. (1.60)

$$A_{CP} = \frac{A_{CP}^{\rm meas}}{(1-2\chi)} \, .$$

The statistical precision on the CP asymmetry measurement is also diluted by the same factor. The time integrated mixing probability is $\chi = 0.181 \pm 0.004$.

4.7 Results

The results of the fit on the full Run1 – Run6 dataset, with the requirement of $E_\gamma > 2.2$ GeV, is reported in Table 4.21. The measured value of CP asymmetry for the inclusive $B \to X_{s+d}\gamma$ decay, including the correction due to the neutral mis-tag fraction is:

$$A_{CP} = -0.12 \pm 0.11 \pm 0.03 \tag{4.6}$$

where the first error is statistical and the second systematic. This is the combined result of inclusive $b \to s\gamma$ and $b \to d\gamma$ decays. Figures 4.33 and 4.34 show the projection of the likelihood on the two PDF's for all the components in B and $\bar{B}$ events, respectively. The corresponding signal enhanced projections are shown in Figures 4.35 and 4.36. They are obtained by applying a cut on the likelihood ratio $R > 0.7$ for each event, where R is the ratio of signal over total events. In order to plot the distibution of $m_{\rm ES}$ (BDT), we first integrate out the $m_{\rm ES}$ (BDT) variable,

Parameter	Fit result
A_{CP} (SIG)	-0.094 ± 0.089
N_{SIG}	300 ± 27
A_{CP} (COMB)	-0.08 ± 0.22
N_{COMB}	-355 ± 125
A_{CP} ($q\overline{q}$)	0.008 ± 0.010
$N_{q\overline{q}}$	13428 ± 166
$q\overline{q}$ $m_{\rm ES}\xi$	-24.9 ± 1.0
$B\bar{B}_p$ $BDT\mu_1$	-0.226 ± 0.014
$B\bar{B}_p$ BDT σ_1	0.0682 ± 0.0094
$B\bar{B}_p$ BDT μ_2	0.0971 ± 0.0055
$B\bar{B}_p$ BDT σ_2	0.1110 ± 0.0039
$B\bar{B}_p$ BDT f	0.1552 ± 0.021
$B\bar{B}_c$ $BDT\mu_1$	0.1744 ± 0.0065
$B\bar{B}_c$ BDT σ_1	0.0565 ± 0.0043
$B\bar{B}_c$ BDT μ_2	-0.2896 ± 0.0059
$B\bar{B}_c$ BDT σ_2	0.0238 ± 0.0042
$B\bar{B}_c$ BDT f	0.554 ± 0.077

Table 4.21. Fit results on the full Run1 – Run6 dataset for events with $E_\gamma > 2.2$ GeV.

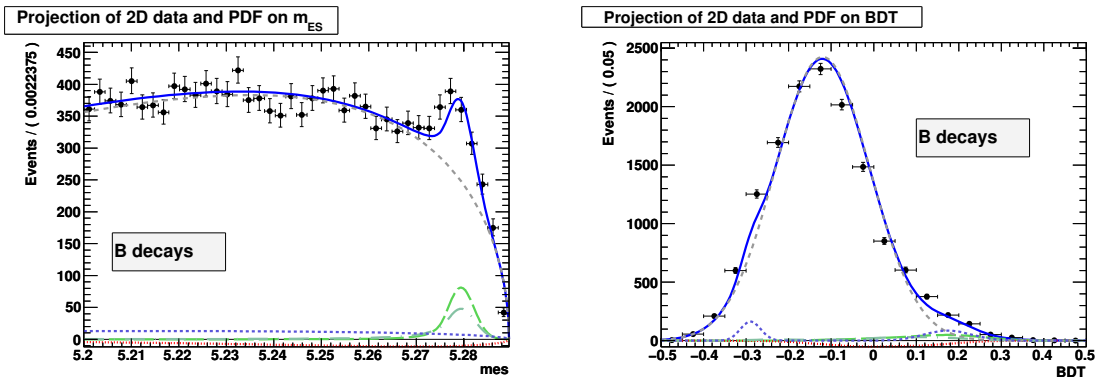

Figure 4.33. Projections of the likelihood on the two PDF's for all the components in the fit, tagged as B decays: long-dashed green curve is signal, dashed grey is continuum, dotted red is combinatoric, dot-dashed dark green and dashed dark blue are the peaking and combinatoric $B\bar{B}$ components, respectively.

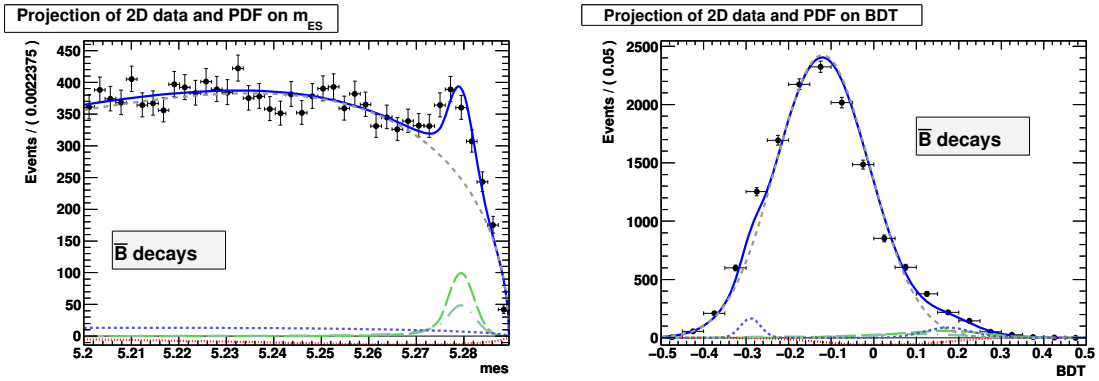

Figure 4.34. Projections of the likelihood on the two PDF's for all the components in the fit, tagged as $\bar{B}$ decays: long-dashed green curve is signal, dashed grey is continuum, dotted red is combinatoric, dot-dashed dark green and dashed dark blue are the peaking and combinatoric $B\bar{B}$ components, respectively.

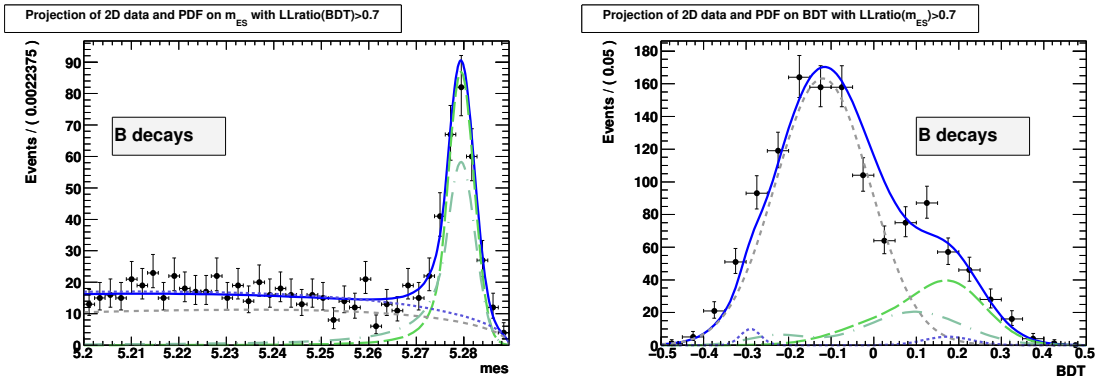

Figure 4.35. Signal enhanced projections of the likelihood on the two PDF's for the following components in the fit, tagged as B decays: long-dashed green curve is signal, dashed grey is continuum, dot-dashed dark green and dashed dark blue are the peaking and combinatoric $B\bar{B}$ components, respectively.

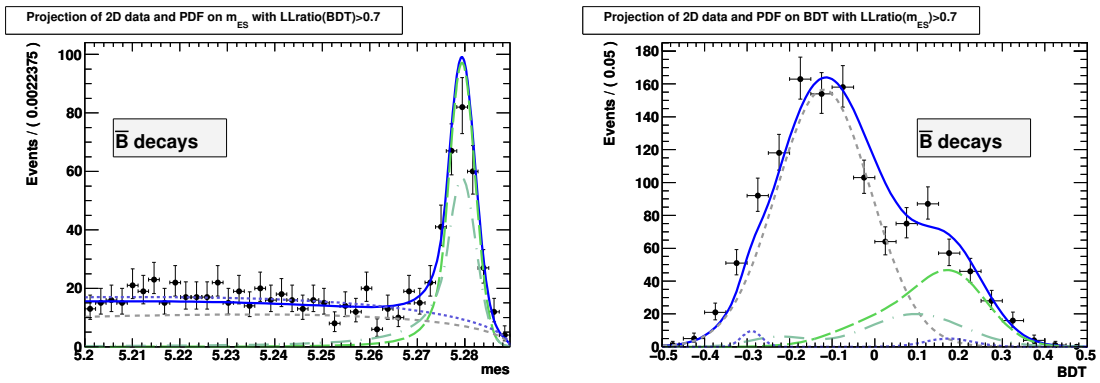

Figure 4.36. Signal enhanced projections of the likelihood on the two PDF's for the following components in the fit, tagged as $\bar{B}$ decays: long-dashed green curve is signal, dashed grey is continuum, dot-dashed dark green and dashed dark blue are the peaking and combinatoric $B\bar{B}$ components, respectively.

then we calculate the ratio R for the remaining BDT ($m_{\rm ES}$) PDF. Figure 4.37 shows the comparison of various previous CP asymmetry measurements performed by the *BABAR* and Belle collaborations, compared to the latest corresponding theoretical calculations. The result of this thesis, obtained over a total integrated luminosity of 427 fb^{-1}, supersedes the previous *BABAR* result [62]: $A_{CP} = 0.10 \pm 0.18 \pm 0.05$ – not shown in Figure 4.37 – which was obtained on a data sample of 210 fb^{-1}.

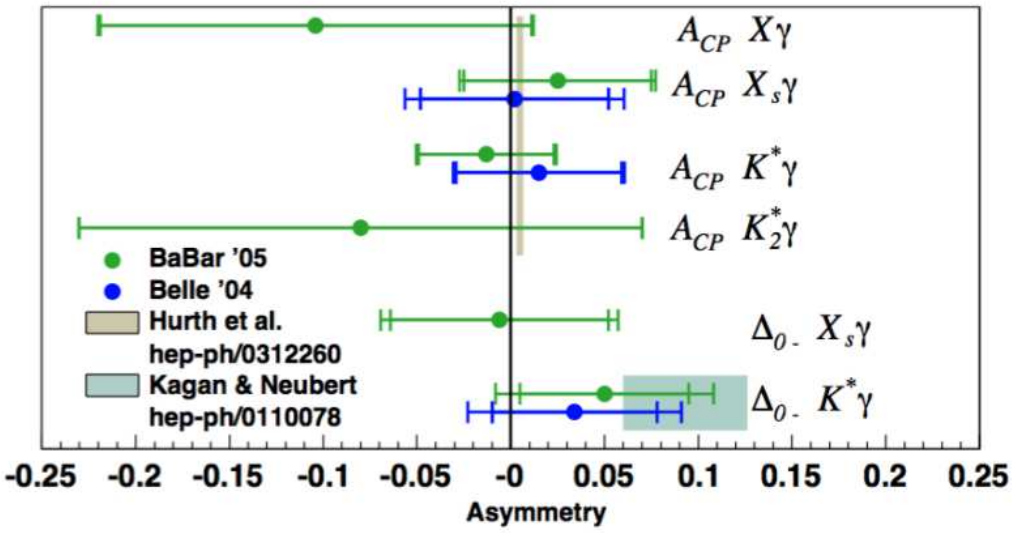

Figure 4.37. Previous *CP* asymmetry measurements by the *BABAR* (green) and Belle (blue) collaborations. The light red line is the corresponding SM expectation, according to Reference [52].

CONCLUSIONS

In this work we presented two processes governed by Flavor-Changing Neutral Currents. Since FCNC's are not yet strongly constrained, they are ground for possible new physics searches. The experimental precision achieved so far is still too large for direct detection of NP effects, but we can still reduce the allowed space of parameters.

The first of the two studies presented here is the search for $B^0 \to \ell^+\ell'^-$ decays, with ℓ, ℓ' being either an electron or a muon. They are very sensitive to NP searches in the scenario of Minimum Flavor Violation, where an eventual shift in the branching fraction is a clear signal of new heavy states in the loops governing these decays. Even if this search is facilitated by the fact that it is relatively easy to find a hard electron or muon in the detector, it is still an experimental challenge because these decays are very rare. Using an optimized event selection and a maximum likelihood fit, we placed upper limits for the branching fractions, at 90% confidence level, of $\mathcal{BR}(B^0 \to e^+e^-) < 11.3 \times 10^{-8}$, $\mathcal{BR}(B^0 \to \mu^+\mu^-) < 5.2 \times 10^{-8}$, and $\mathcal{BR}(B^0 \to e^\pm\mu^\mp) < 9.2 \times 10^{-8}$, which are still the world smallest values in the electron channels.

The second study is the measurement of *CP* asymmetry in inclusive $B \to X_s\gamma$ decays. This class of decays is one of the richest in flavor physics, because of the wide variety of measurements that they offer. Among the possibilities, we chose the measurement of *CP* asymmetry, since all the previous works have focused on other measurements. This is an inclusive study, which makes it experimentally very different from the $B^0 \to \ell^+\ell'^-$ work. It is particularly challenging because of the rare nature of these decays and because of the very high background associated with the sole reconstruction of a photon. In order to reduce the backgrounds, we used a recoil technique, in which one of the two B's in the event is reconstructed in very

well known hadronic modes. However, the associated efficiency is very small, thus requiring a large sample of B decays, which can be obtained only at a B Factory. We used a multivariate discriminant in order to improve the signal selection, and extracted the CP asymmetry from a maximum likelihood fit.

The result $A_{CP}(B \to X_{s+d}\gamma) = -0.12 \pm 0.11$ (stat.) ± 0.03 (syst.) is an improvement with respect to the latest measurement with the same technique.

APPENDIX A

$B^0 \to \ell^+\ell'^-$ STUDIES

Variable	$\pi^+\pi^-$MC	$\pi^+\pi^-$ Data	Correction
SIG $m_{\rm ES}$ μ	5.27960 ± 0.00002	5.27940 ± 0.00007	-0.00020 ± 0.00007
SIG $m_{\rm ES}$ σ_L	0.00253 ± 0.00002	0.00227 ± 0.00008	0.895 ± 0.037
SIG $m_{\rm ES}$ σ_R	0.00248 ± 0.00001	0.00229 ± 0.00014	0.923 ± 0.062
SIG $m_{\rm ES}$ α_L	0.1014 ± 0.0011	0.189 ± 0.040	1.87 ± 0.40
SIG $m_{\rm ES}$ σ_R	0.	0.041 ± 0.036	–
SIG ΔE μ	-0.00135 ± 0.00030	-0.0016 ± 0.0018	-0.0003 ± 0.0018
SIG ΔE σ_L	0.0298 ± 0.0014	0.02673 ± 0.00076	0.91 ± 0.050
SIG ΔE σ_R	0.02531 ± 0.00095	0.03109 ± 0.00040	0.76 ± 0.050
SIG ΔE α_L	0.0738 ± 0.043	0.112 ± 0.020	1.73 ± 0.92
SIG ΔE α_R	0.0068 ± 0.0055	0.0570 ± 0.0099	0.050 ± 0.011
SIG $\mathcal{F}$ μ	-0.0044 ± 0.0044	0.026 ± 0.039	0.030 ± 0.039
SIG $\mathcal{F}$ σ_L	0.6832 ± 0.0030	0.716 ± 0.022	1.048 ± 0.033
SIG $\mathcal{F}$ σ_R	0.3835 ± 0.0025	0.423 ± 0.030	1.103 ± 0.079

Table A.1. Summary of the fit parameters, which are obtained with 290k events of Monte Carlo samples of the two hadronic decays considered in the fits.

Data vs. Monte Carlo comparison

In order to quantify the agreement between data and Monte Carlo in describing the shapes of signal Monte Carlo events, we use the abundance of $B \to \pi\pi$ events in the hh sample.

Fixing all the parameters of the fit of this sample to the results of Section 3.6.1.1, we float one by one the shapes of the three variables we are interested in ($m_{\rm ES}$, ΔE, and Fisher) and we compare that to the values for the same parameters, determined from Monte Carlo. In the case of the fit for ΔE, we extended the fit range from $[-100, 150]$ MeV to $[-150, 150]$ MeV, to be sensitive to the left-side tail.

The result of this study, given in Table A.1, allows to calculate correction factors to the parameterizations of signal PDFss for the three leptonic modes. In particular, the mean values are corrected for the difference between the corresponding value for $B \to \pi\pi$ data and Monte Carlo, while the other parameters are scaled according to the ratio. The corrections are quoted in the last column of Table A.1. The errors on the scaling factors are computed by propagating the errors.

PID selection criteria

We briefly describe the selection criteria for the identification of muons and electrons. The selection of muons relies on the IFR and is obtained with a neural network trained with the following variables:

- energy released in the calorimeter by the muon ($E_{\rm cal}$);
- the number of IFR layers hit by a track and used for the construction of a three-dimensional IFR cluster (N_L);
- the number of interaction lengths (λ_{meas}) crossed by the track in the detector. It is estimated by extrapolating the track to the last IFR layer;
- the difference between the measured and the expected interaction lengths ($\Delta\lambda$);
- the χ^2/d.o.f. of the fit of the track extrapolation in the IFR;
- the track continuity, defined as:

$$T_c = \frac{N_L}{L_h - F_h}, \tag{A.1}$$

  where F_h and L_h are the first and last layer hit by a track;
- the average multiplicity of IFR strips hit by a track ($\bar{m}$) and its standard deviation (σ_m).

The control sample used is composed of muons coming from the following process:

$$e^+e^- \to \mu^+\mu^-\gamma$$

The pion control sample, for studies of mis-identification probability of muons into pions, is:

$$\tau^+(\to \ell\nu_\ell\nu_\tau), \tau^-(\to \pi\pi\pi\nu_\tau)$$

The choice of these selection variables has been made by training the neural network adding one variable at a time until the best configuration was found.

The selection of electrons relies DCH, DIRC and EMC. Using samples of electrons, pions, kaons and protons, the electron identification is performed in two steps. The first is to apply loose selection criteria on electrons coming from Bhabha scattering events; on pions coming from $K^0_S \to \pi\pi$ events; on kaons coming from $D^* \to D^0\pi,\ D^0 \to K\pi$; on protons coming from $\Lambda \to p\pi$. The second step is to build probability density function for each discriminating variable. Under the assumption of independent measurements from the individual sub-detectors, they are combined to compute the likelihood $L(\xi)$ for each particle hypothesis, $\xi \in \{e;\pi;K;p\}$. Weighting the individual likelihoods with *a priori* probabilities p_ξ, the likelihood fraction f_L is computed as:

$$f_L = \frac{p_e L(e)}{p_e L(e) + p_\pi L(\pi) + p_K L(K) + p_p L(p)}$$

A track is then selected as an electron if it passes the preselection cuts and a given cut on f_L, which may vary between 0 and 1, chosen to be $f_L > 0.95$. Finally, electron identification efficiency and hadron fake rates are measured on data.

Fit of on-resonance sideband sample

Once the yields of $q\bar{q}$ components are determined by the fit of the hh data sample (see Section 3.6.1), we can determine the expected background yields in the leptonic samples by knowing the relative fractions of continuum events in each PID sample.

This information can be extracted from a fit of the sideband of the on-resonance sample, defined by $m_{\rm ES} < 5.25$ GeV/c^2 and $\Delta E > -50$ MeV. The variables used in the fit are the same as in Section 3.6.1, except $m_{\rm ES}$, which has been removed after the tight selection cut. The likelihood for each PID class is written as:

$$
\begin{aligned}
\mathcal{L} \quad = \quad & \frac{e^{-(\epsilon^{PID}_{b\pi\pi}\times N_{b\pi\pi}+\epsilon^{PID}_{bK\pi}\times N_{bK\pi}+\epsilon^{PID}_{bKK}\times N_{bKK})/N^{PID}}}{N^{PID}\sqrt{(\epsilon^{PID}_{b\pi\pi}\times N_{b\pi\pi}+\epsilon^{PID}_{bK\pi}\times N_{bK\pi}+\epsilon^{PID}_{bKK}\times N_{bKK})!}}\cdot\prod_{i=1}^{N}\Big\{ \\
N_{b\pi\pi}\cdot \quad & \epsilon^{PID}_{b\pi\pi}\cdot P_{bhh}(\Delta E)_i\cdot P^{+}_{\pi}(\theta^{+}_{c})_i\cdot P^{-}_{\pi}(\theta^{-}_{c})_i\cdot P_{bhh}(\mathcal{F})_i + \\
N_{bK\pi}\cdot \quad & \frac{1-A_{bK\pi}}{2}\cdot\epsilon^{PID}_{bK\pi}\cdot P_{bhh}(\Delta E)_i\cdot P^{+}_{K}(\theta^{+}_{c})_i\cdot P^{-}_{\pi}(\theta^{-}_{c})_i\cdot P_{bhh}(\mathcal{F})_i + \\
N_{bK\pi}\cdot \quad & \frac{1+A_{bK\pi}}{2}\cdot\epsilon^{PID}_{bK\pi}\cdot P_{bhh}(\Delta E)_i\cdot P^{+}_{\pi}(\theta^{+}_{c})_i\cdot P^{-}_{K}(\theta^{-}_{c})_i\cdot P_{bhh}(\mathcal{F})_i + \\
N_{bKK}\cdot \quad & \epsilon^{PID}_{bKK}\cdot P_{bhh}(\Delta E)_i\cdot P^{+}_{K}(\theta^{+}_{c})_i\cdot P^{-}_{K}(\theta^{-}_{c})_i\cdot P_{bhh}(\mathcal{F})_i\Big\}, \qquad \text{(A.2)}
\end{aligned}
$$

where $N_{bK\pi}$, N_{bKK}, and $N_{b\pi\pi}$ are the total yields (*i.e.* the sum on the four PID classes) of $q\bar{q}$ background yields which the DIRC PDF identifies as $K\pi$, KK or whatever is left ($\pi\pi$ and ll); N^{PID} is the total number of events in the considered PID category; the P functions are those previously defined (see Eq. 4.3). The fit is performed simultaneously on the four PID categories ($PID = hh, 2\mu, 1e1\mu$, and $2e$) and the fraction ϵ^{PID}_{BKG} represents the fraction of events of each background component BKG ($BKG = b\pi\pi, bK\pi$, and bKK) in the class PID (that we want to determine in this fit). The simultaneous fit forces by construction the four fractions to preserve the normalization, imposing $\epsilon^{hh}_{bhh} = 1 - \sum_{PID\neq hh}\epsilon^{PID}_{bhh}$.

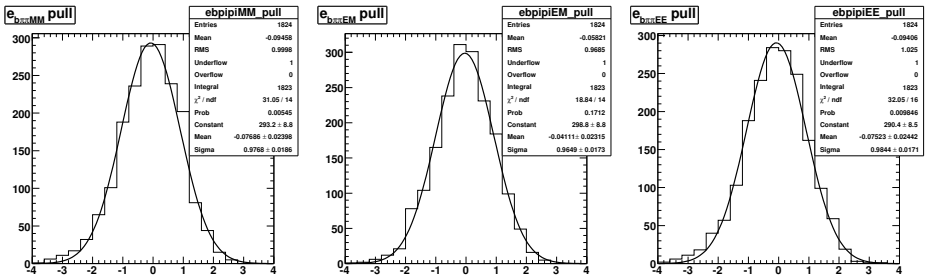

Figure A.1. Distribution of the pull of the three non-zero floated split fractions, i.e. of the $b\pi\pi$ component into the three leptonic categories, obtained from a set of toy Monte Carlo experiments performed with the expected input values Run1-5 on-resonance sideband dataset.

Fit Validation

We validate this fit by performing a set of toy Monte Carlo experiments, which allow to check the presence of any bias on the split fractions. The pull and the error distributions for the three non-zero floated split fractions are given in Figure A.1. We can see that the mean values of the pulls are different from zero, so we estimate the bias by multiplying these means by the error of the fits, obtaining:

- $\epsilon^{2\mu}_{b\pi\pi}$: the bias is $-1.8 \cdot 10^{-5}$, compared to the result of the fit $(1.55 \pm 0.24)10^{-3}$;
- $\epsilon^{1\mu 1e}_{b\pi\pi}$: the bias is $-1.1 \cdot 10^{-5}$, compared to the result of the fit $(1.88 \pm 0.26)10^{-3}$;
- $\epsilon^{2e}_{b\pi\pi}$: the bias is $-1.8 \cdot 10^{-5}$, compared to the result of the fit $(1.55 \pm 0.24) \cdot 10^{-3}$.

Fit on Run1-5 on-resonance sideband data

The output of the fit on the on-resonance sideband is given in Table A.2.

The values of the $q\bar{q}$ split fractions are used to calculate the $q\bar{q}$ background in the nominal fit of leptonic yields, according to this formula:

$$N^{ll}_{b\pi\pi} = \frac{N^{hh}_{b\pi\pi}}{\epsilon^{hh}_{b\pi\pi}} \cdot \epsilon^{ll}_{b\pi\pi} \tag{A.3}$$

where:

Yield	Value of the yield
$N_{b\pi\pi}$	27707 ± 176
$N_{bK\pi}$	18601 ± 155
$A_{CP}^{bK\pi}$	0.0054 ± 0.0082
N_{bKK}	12172 ± 119
$\epsilon_{b\pi\pi}^{2\mu}$	0.00155 ± 0.00024
$\epsilon_{b\pi\pi}^{1\mu 1e}$	0.00188 ± 0.00026
$\epsilon_{b\pi\pi}^{2e}$	0.00155 ± 0.00024
$\epsilon_{bK\pi}^{2\mu}$	$0. \pm 0.000036$
$\epsilon_{bK\pi}^{1\mu 1e}$	$0. \pm 0.00027$
$\epsilon_{bK\pi}^{2e}$	$0. \pm 0.000028$
$\epsilon_{bKK}^{2\mu}$	$0. \pm 0.000041$
$\epsilon_{bKK}^{1\mu 1e}$	$0. \pm 0.000047$
ϵ_{bKK}^{2e}	$0. \pm 0.000041$

Table A.2. Result of the fit of the sideband of the on-resonance Run1-5 dataset to obtain the splitting fractions for the $q\bar{q}$ categories of the leptonic decays.

- $N_{b\pi\pi}^{ll}$ is the number of expected background events in each leptonic category;
- $N_{b\pi\pi}^{hh}$ is the number of background events in the hh category, obtained from the fit to the hh dataset (i.e. the $N_{b\pi\pi}$ of the fit of Table 3.13).
- $\epsilon_{b\pi\pi}^{ll}$ is the splitting fraction to the ll category;
- $\epsilon_{b\pi\pi}^{hh}$ is the splitting fraction to the hh category, which is calculated as:

$$\epsilon_{b\pi\pi}^{hh} = 1 - \sum_{LL} \epsilon_{b\pi\pi}^{LL} \tag{A.4}$$

  where the index LL runs over the three leptonic categories.

In order to take into account the bias of the fit to the splitting fractions and the bias of the fit to the hh category, Eq. A.3 actually reads:

$$N_{b\pi\pi}^{ll} = \frac{N_{b\pi\pi}^{hh} + BIAS(N_{b\pi\pi}^{hh}) \cdot \sigma(N_{b\pi\pi}^{hh})}{\epsilon_{b\pi\pi}^{hh} + BIAS(\epsilon_{b\pi\pi}^{hh}) \cdot \sigma(\epsilon_{b\pi\pi}^{hh})} \cdot \left[\epsilon_{b\pi\pi}^{ll} + BIAS(\epsilon_{b\pi\pi}^{ll}) \cdot \sigma(\epsilon_{b\pi\pi}^{ll})\right] \tag{A.5}$$

where:

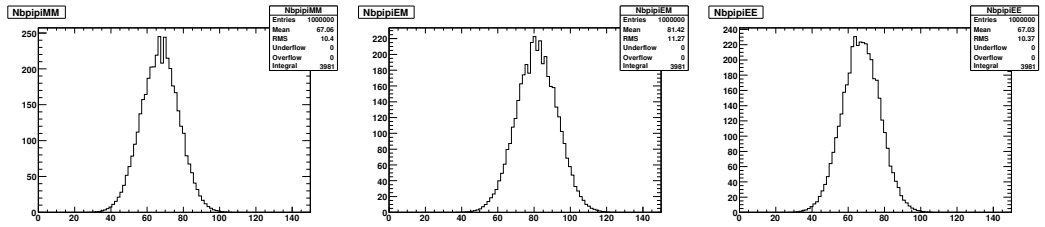

Figure A.2. Expected number of background yields for the three leptonic categories, 2μ (left), $1\mu 1e$ (center) and $2e$ (right), estimated with 1 million toy Monte Carlo experiments, which includes all the biases to the fits and the systematic error associated to the number of background events in the category hh. Each event in these plots is weighted by the correlation between all variables used in the calculations.

- the three $BIAS$es are the means of the pull distributions obtained from the toy Monte Carlo experiments;
- $\sigma(N_{b\pi\pi}^{hh})$ includes the statistical error on the background yield calculated from the fit to hh;
- $\sigma(\epsilon_{b\pi\pi}^{PID})$ are the statistical errors on the fractions obtained from the fit.

We use a Monte Carlo technique to calculate the total error on the number of background events, taking into account all the existing correlations. We generate one million events, extracting a random number according to a multidimensional Gaussian centered around the quantity $BIAS \cdot \sigma$ for each of the three variables in the formula. The results of the toy Monte Carlo experiments are shown in Figure A.2.

In principle, we should also take into account the correlations between all the variables used in the calculation of $N_{b\pi\pi}^{ll}$ and use them as a weight for each event in the distributions. Anyhow we found all the correlations to be negligible in the fit, so that this feature is neglected when performing the extraction.

Table A.3 shows the biases and the expected number of background events in each of the three leptonic categories. The expected numbers of background events of Table A.3 are compatible with the blind fit reported in Table 3.22.

Category	Fraction	Bias of the Fraction	Expected number of background events
2μ	0.00155 ± 0.00024	0.07686 ± 0.02398	67.06 ± 10.40
$1\mu 1e$	0.00188 ± 0.00026	0.04111 ± 0.02315	81.42 ± 11.27
$2e$	0.00155 ± 0.00024	0.07523 ± 0.02442	67.03 ± 10.37

Table A.3. Relative fractions of continuum events in the three leptonic samples, associated biases and numbers of expected background events in the three leptonic categories. Each bias is the mean of the corresponding pull distribution obtained with toy Monte Carlo experiments.

parameter	$\sigma_+(N_{\mu\mu})$	$\sigma_-(N_{\mu\mu})$	$\sigma_+(N_{e\mu})$	$\sigma_-(N_{e\mu})$	$\sigma_+(N_{ee})$	$\sigma_-(N_{ee})$
$m_{\rm ES}$ ξ	0.039	-0.039	0.012	-0.012	0.028	-0.028
ΔE p_1	0.010	-0.010	0.001	-0.001	0.018	-0.019
ΔE p_2	0.039	-0.040	0.003	-0.003	0.003	-0.003
Fisher μ_1	0.026	-0.025	0.012	-0.012	0.028	-0.028
Fisher μ_2	0.010	0.003	0.030	-0.031	0.240	-0.243
Fisher σ_1	0.010	-0.032	0.019	-0.019	0.049	-0.050
Fisher σ_2	0.003	-0.001	0.009	-0.010	0.220	-0.246
Fisher f	0.014	0.007	0.069	-0.076	0.415	-0.399
Total	0.065	-0.070	0.080	-0.087	0.323	0.532

Table A.4. Summary of systematic errors on the leptonic yields, induced by the knowledge of background PDF parameters.

Systematics Error induced by $q\overline{q}$ PDF's

For each parameter of continuum background PDF's that is fixed in the fit of the leptonic samples, we evaluate the systematic effect by shifting the value of the parameter by $\pm 1\sigma$ from its nominal value and using the observed difference on the leptonic yield as a systematic. We use the error returned by the fit to the hh sample as σ. To be conservative, all the errors are taken as uncorrelated. A summary of the contributions is given in Table A.4.

Systematics Error induced by signal PDF's

For each parameter of signal PDF's that is fixed in the fit of the leptonic samples, we evaluate the systematic effect by shifting the value of the parameter by $\pm 1\sigma$ from the value obtained correcting signal Monte Carlo parameters according to the study of Appendix A and by using the observed difference on the leptonic yield as a systematic. Since these effects are dominated by the statistical precision on the fit of the hh sample, we quote the largest contribution as the systematic error.

parameter	$\sigma_+(N_{\mu\mu})$	$\sigma_-(N_{\mu\mu})$	$\sigma_+(N_{e\mu})$	$\sigma_-(N_{e\mu})$	$\sigma_+(N_{ee})$	$\sigma_-(N_{ee})$
$m_{\rm ES}$ μ	0.224	-0.233	0.006	-0.089	0.145	-0.148
$m_{\rm ES}$ σ_L	0.380	-0.450	0.049	-0.049	0.148	-0.457
$m_{\rm ES}$ σ_R	0.211	-0.208	0.014	-0.035	0.135	-0.127
$m_{\rm ES}$ α_L	0.321	-0.127	0.025	-0.026	1.065	-0.913
$m_{\rm ES}$ α_R	0.080	-0.089	0.030	-0.012	0.041	-0.048
ΔE μ	0.323	-0.371	0.019	-0.069	0.036	-0.031
ΔE σ_L	0.010	-0.017	0.059	-0.009	0.098	-0.104
ΔE σ_R	0.048	-0.048	0.069	-0.042	0.001	-0.093
ΔE α_L	0.057	-0.062	0.146	-0.163	0.042	-0.067
ΔE α_R	0.047	-0.048	0.013	-0.013	0.042	-0.006
Fisher μ	0.144	-0.208	0.023	-0.037	0.322	-0.354
Fisher σ_L	0.128	-0.208	0.001	-0.002	0.252	-0.274
Fisher σ_R	0.146	-0.153	0.008	-0.013	0.102	-0.107
Total	0.380	-0.455	0.166	-0.184	1.190	-1.056

Table A.5. Summary of systematic errors on the leptonic yields, induced by the knowledge of signal PDF parameters.

APPENDIX B

RADIATIVE B DECAYS

Extrapolation of the photon spectrum

The theoretical error on the total inclusive $B \to X_s\gamma$ branching ratio is of the order of 10%. Unfortunately most of the theoretical uncertainty in an inclusive branching fraction measurement derives from other sources. In order to isolate the signal region from the large background (1000 times bigger) a cut on the phase space region is applied. The extrapolation to the full phase space may introduce very large uncertainties and model dependence. Only the high part of the $B \to X_s\gamma$ photon spectrum is accessible from an experimental point of view. Therefore only the branching ratio for $B \to X_s\gamma$ with $E_\gamma > E_\gamma^{\rm min}$ could be directly measured. To obtain the total branching ratio, one has to know the fraction R of the $B \to X_s\gamma$ events with $E_\gamma > E_\gamma^{\rm min}$.

The uncertainty on this fraction R is regarded as a *theoretical* uncertainty. The photon energy spectrum cannot be calculated directly using the heavy mass expansion, because the operator product expansion breaks down in the high-energy part of the spectrum, where $E_\gamma \approx m_b/2$.

The fraction R was calculated, for the first time, in [65] using a phenomenological model [66], where the motion of the b quark in the B meson is characterized by two parameters, the average momentum p_F of the b quark and the average mass m_q of the spectator quark.

A theoretical analysis of the problem was presented in [47]. The residual motion of the b quark inside the B meson caused by its soft interactions with the light constituents leads to a modification of the photon energy spectrum, the so called "Fermi motion". Given that the OPE breaks down near the endpoint of the photon energy spectrum, the Fermi motion is included in the heavy-quark expansion by resumming an infinite set of leading-twist corrections into a shape function $F(k_+)$, which governs the light-cone momentum distribution of the heavy quark inside the B meson defined in Eq. [35].

The shape function is a universal, *i.e.* process-independent characteristic of the B meson governing the inclusive decay spectra in processes with massless partons in the final state, such as $B \to X_s\gamma$ and $B \to X_u\,\ell\,\nu$. It is important to note that this function does not describe in an accurate way the distributions in decays into massive partons such as $B \to X_c\,\ell\,\nu$. Therefore, the shape function cannot be determined using the lepton spectrum in semileptonic decays of B mesons, for which high-precision data exist.

On the other hand, there is some useful theoretical information on the moments of the shape function, which are related to the forward matrix elements of local operators:

$$A_n = \int \mathrm{d}k_+\, k_+^n\, F(k_+) = \frac{1}{2m_B}\,\langle B|\,\bar{b}\,(iD_+)^n b\,|B\rangle\,. \tag{B.1}$$

The first three moments satisfy $A_0 = 1$, $A_1 = 0$ and $A_2 = \frac{1}{3}\mu_\pi^2$, where $\mu_\pi^2 = -\lambda_1$ is related to the kinetic energy of the b quark inside the B meson (Section 1.4). The condition $A_1 = 0$, which is a consequence of the equations of motion, ensures that the quark mass m_b entering the theoretical expressions is the pole mass.

Let $P_{\rm p}(y_{\rm p})$ be the photon energy spectrum in the parton model, where $y_{\rm p} = 2E_\gamma/m_b$ with $0 \le y_{\rm p} \le 1$. The result of including the effects of Fermi motion and calculating the physical spectrum $P(y)$ as a function of the variable $y = 2E_\gamma/m_B$, to the leading-twist approximation,is given by the convolution [35]

$$P(y)\,\mathrm{d}y = \int \mathrm{d}k_+\, F(k_+)\,\Big[P_{\rm p}(y_{\rm p})\,\mathrm{d}y_{\rm p}\Big]_{y_{\rm p}=y_{\rm p}(k_+)}\,, \tag{B.2}$$

where $y_{\rm p}(k_+)$ is obtained by replacing m_b in the definition of $y_{\rm p}$ with the "effective mass" $m_b^* = m_b + k_+$, *i.e.* $y_{\rm p}(k_+) = 2E_\gamma/m_b^* = ym_B/m_b^*$. Because the support of the shape function is restricted to the range $-m_b \le k_+ \le m_B - m_b$, it follows that $0 \le y \le 1$. Denoting $B_{\rm p}(\delta_{\rm p})$ the integrated branching ratio calculated in the parton model, which is given by an integral over the spectrum $P_{\rm p}(y_{\rm p})$ with a cutoff $\delta_{\rm p}$ defined

by the condition that $E_\gamma \geq \frac{1}{2}(1-\delta_{\rm p})m_b$, from Eq. B.2, follows that the corresponding physical quantity $B(\delta)$ is given by:

$$B(\delta) = \int\limits_{m_B(1-\delta)-m_b}^{m_B-m_b} \mathrm{d}k_+ \, F(k_+) \, B_{\rm p}\left(1 - \frac{m_B(1-\delta)}{m_b + k_+}\right) . \tag{B.3}$$

This relation is such that $B(1) = B_{\rm p}(1)$, implying that the total branching ratio is not affected by Fermi motion; indeed, the $1/m_Q^2$ corrections are the only power corrections to the total branching ratio.

A simple ansatz for the distribution function is:

$$F(k_+) = N(1-x)^a e^{(1+a)x}; \quad x = \frac{k_+}{\overline{\Lambda}} \leq 1 \quad \textit{and} \quad a = f(\overline{\Lambda}, \lambda_1), \tag{B.4}$$

where $\bar{\Lambda} = m_B - m_b$. The parameters N, a are chosen such that the first three moments of $F(k_+)$ satisfy the relations mentioned after Eq. B.1. The parameter a can be related to the second moment, yielding $A_2 = \frac{1}{3}\mu_\pi^2 = \bar{\Lambda}^2/(1+a)$. Thus, the b-quark mass (or $\bar{\Lambda}$) and the quantity μ_π^2 (or a) are the two parameters of this function.

For a graphical illustration of the sensitivity of the results to the parameters of the shape function, the predictions for the Standard Model branching ratio as a function of the energy cutoff $E_\gamma^{\rm min}$ are shown in Figure B.1. The gray line shows the result obtained using the same parameters as for the solid line, but with a Gaussian ansatz $F(k_+) = N\,(1-x)^a e^{-b(1-x)^2}$ for the shape function.

Comparing the two upper plots in Figure B.1, it's clear that the uncertainty due to the value of the b-quark mass is the dominant one. Variations of the parameter μ_π^2 have a much smaller effect on the partially integrated branching ratio, and also the sensitivity to the functional form adopted for the shape function turns out to be small. This behaviour is a consequence of global quark–hadron duality, which

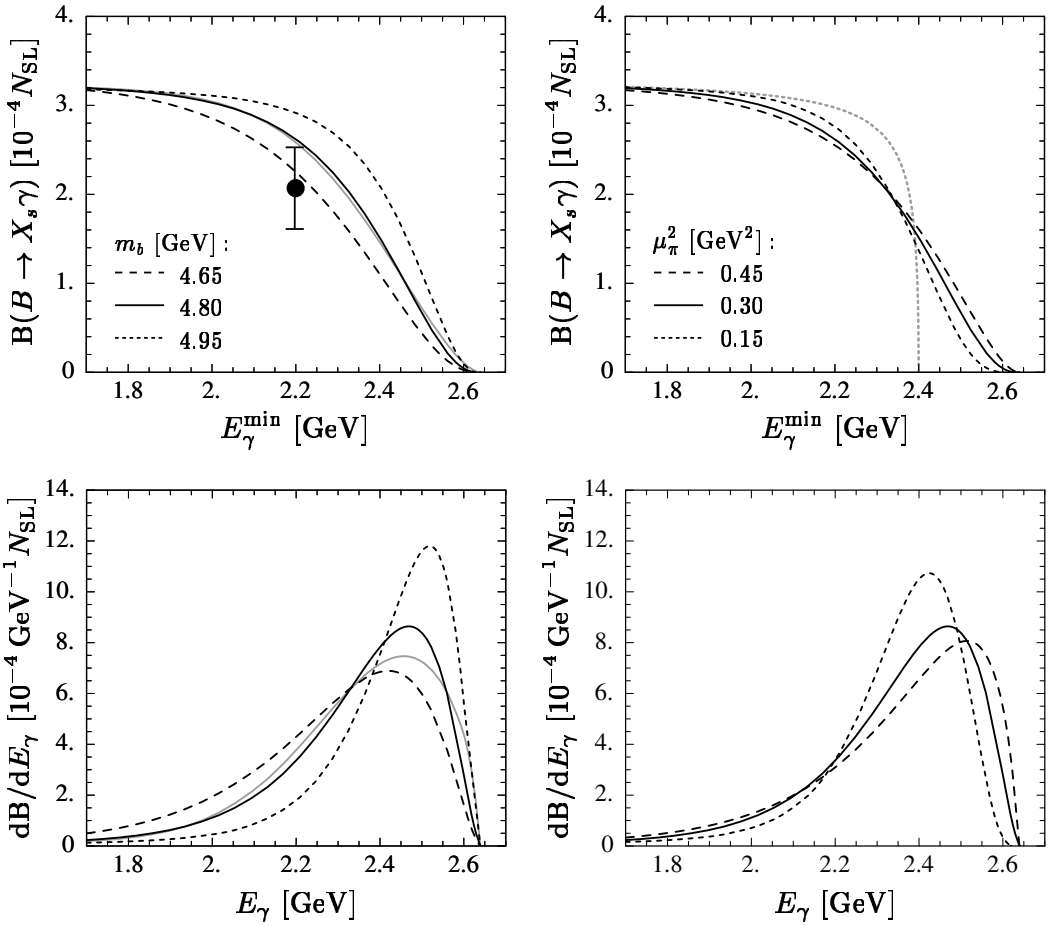

Figure B.1. The model dependence of the E_γ spectrum in the B meson rest frame, taken from the paper by Kagan and Neubert [47]. The spectra are shown for different choices of quark mass and Fermi momentum. Also shown are the integrals of the spectra as a function of the lower bound of integration, E_γ^{min}. The data point represents the first CLEO measurement, as provided to Kagan and Neubert by private communication.

ensures that even partially integrated quantities are rather insensitive to bound-state effects. The strong remaining dependence on the b-quark mass is simply due to the transformation by Fermi motion of phase-space boundaries from parton to hadron kinematics.

The spread of results obtained by varying m_b between 4.65 and 4.95 GeV/c^2 (with μ_π^2 adjusted as described above) represents the amount of model dependence resulting from the inclusion of Fermi motion.

An important observation is that the shape of the photon spectrum is practically insensitive to physics beyond the SM. A precise measurement of the photon spectrum allows to determine the parameters of the shape function.

The latter information is an important input for the determination of the CKM matrix element V_{ub}. One takes advantage of the universality of the shape function to lowest order in Λ_{QCD}/m_b. The same shape function occurs in the description of non-perturbative effects in the endpoint region of the $B \to X_s\gamma$ photon spectrum and of the $B \to X_u\ell\nu$ charged-lepton spectrum up to higher $1/m_b$ corrections. Thus, from the photon spectrum one can determine the shape function; with the help of the latter and of the measurement of the charged-lepton spectrum of $B \to X_u\ell\nu$, one can extract a value for V_{ub}.

Systematics Error induced by the fit PDF's

For each parameter of the PDF's that is fixed in the fit, we evaluate the systematic effect by shifting the value of the parameter by $\pm 1\sigma$ from its nominal value and using the observed difference on the signal *CP* asymmetry as a systematic. To be conservative, all the errors are taken as uncorrelated. A summary of the contributions is given in Tables B.1, B.2, B.3 and B.4 for signal, combinatoric, continuum and $B\bar{B}$ parameters, respectively.

parameter	$\sigma_+(A_{CP})$	$\sigma_-(A_{CP})$
$m_{\rm ES}\mu$	0.	0.
$m_{\rm ES}\sigma$	-0.0021	0.0038
$m_{\rm ES}\alpha$	-0.0029	0.0001
$m_{\rm ES}$ N	-0.0028	-0.0030
BDT μ_1	0.0025	-0.0025
BDT μ_2	0.0036	-0.0016
BDT σ_1	0.0018	-0.0033
BDT σ_2	0.0031	-0.0026
BDT f	0.0045	-0.0005
Total	0.0085	-0.0071

Table B.1. Summary of systematic errors on the signal *CP* asymmetry, induced by the knowledge of signal PDF parameters.

parameter	$\sigma_+(A_{CP})$	$\sigma_-(A_{CP})$
$m_{\rm ES}\xi$	0.0032	-0.0019
BDT μ_1	0.0025	-0.0025
BDT μ_2	0.0009	-0.0037
BDT σ_1	0.0016	-0.0028
BDT σ_2	0.0022	-0.0038
BDT f	0.0022	-0.0011
Total	0.0054	-0.0069

Table B.2. Summary of systematic errors on the signal *CP* asymmetry, induced by the knowledge of combinatoric background PDF parameters.

parameter	$\sigma_+(A_{CP})$	$\sigma_-(A_{CP})$
BDT μ_1	0.	0.
BDT μ_2	0.0026	-0.0026
BDT σ_1	0.0036	-0.0006
BDT σ_2	0.0017	-0.0016
BDT f	0.0025	-0.0026
Total	0.0054	-0.0041

Table B.3. Summary of systematic errors on the signal *CP* asymmetry, induced by the knowledge of $q\overline{q}$ background PDF parameters.

parameter	$\sigma_+(A_{CP})$	$\sigma_-(A_{CP})$
$m_{\rm ES}\mu$	0.0035	-0.0035
$m_{\rm ES}\sigma$	0.0038	-0.0012
$m_{\rm ES}\alpha$	0.0003	-0.0051
$m_{\rm ES}$ N	-0.0022	-0.0192
BDT μ_1	0.0029	-0.0026
BDT μ_2	0.0025	-0.0025
BDT σ_1	0.0026	-0.0026
BDT σ_2	0.0025	-0.0026
BDT f	0.0026	-0.0029
Total	0.0081	-0.0211

Table B.4. Summary of systematic errors on the signal *CP* asymmetry, induced by the knowledge of the peaking component of the $B\bar{B}$ background PDF parameters.

BIBLIOGRAPHY

[1] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **86**, 2515 (2001); Belle Collaboration, A. Abashian *et al.*, Phys. Rev. Lett. **86**, 2509 (2001).

[2] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **93**, 131801 (2004); Belle Collaboration, Y. Chao *et al.*, Phys. Rev. Lett. **93**, 191802 (2004).

[3] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. D **71** (2005) 091102 [arXiv:hep-ex/0502019]; K. Abe *et al.* [Belle Collaboration], arXiv:hep-ex/0507037; B. Aubert *et al.* [BABAR Collaboration], arXiv:hep-ex/0507087; B. Aubert *et al.* [BABAR Collaboration], arXiv:hep-ex/0408095; B. Aubert *et al.* [BABAR Collaboration], Phys. Rev. D **71** (2005) 111102 [arXiv:hep-ex/0503011]; B. Aubert *et al.* [BABAR Collaboration], arXiv:hep-ex/0508017; B. Aubert *et al.* [BABAR Collaboration], arXiv:hep-ex/0603040; B. Aubert *et al.* [BABAR Collaboration], arXiv:hep-ex/0507016; B. Aubert *et al.* [BABAR Collaboration], arXiv:hep-ex/0507052;

[4] Y. Grossman and M. P. Worah, Phys. Lett. B **395** (1997) 241 [arXiv:hep-ph/9612269]; M. Ciuchini, E. Franco, G. Martinelli, A. Masiero and L. Silvestrini, Phys. Rev. Lett. **79** (1997) 978 [arXiv:hep-ph/9704274]; Y. Grossman, G. Isidori and M. P. Worah, Phys. Rev. D **58** (1998) 057504 [arXiv:hep-ph/9708305];

[5] A. J. Buras, P. Gambino, M. Gorbahn, S. Jager and L. Silvestrini, "Universal unitarity triangle and physics beyond the standard model", Phys. Lett. B **500**, 161 (2001) [arXiv:hep-ph/0007085];
G. D'Ambrosio, G. F. Giudice, G. Isidori and A. Strumia, "Minimal flavour violation: An effective field theory approach", Nucl. Phys. B **645**, 155 (2002) [arXiv:hep-ph/0207036].

[6] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967);
A. Salam, "Elementary particle theory: relativistic groups and analyticity, Nobel Symp. No. 8, ed. N. Svartholm (Almqvist and Wiksell, Stockholm), 367 (1968).

[7] G. 't Hooft, "Renormalization Of Massless Yang-Mills Fields", Nucl. Phys. B **33** (1971) 173;
G. 't Hooft, "Renormalizable Lagrangians for Massive Yang-Mills Fileds", Nucl. Phys. B **35** (1971) 167.

[8] N. Cabibbo, "Unitary Symmetry and Leptonic Decays", Phys. Rev. Lett. **10**, 531 (1963);
M. Kobayashi and T. Maskawa, "CP Violation In The Renormalizable Theory Of Weak Interaction", Prog. Theor. Phys. **49**, 652 (1973).

[9] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).

[10] M. Neubert, "Effective field theory and heavy quark physics", arXiv:hep-ph/0512222.

[11] G. Buchalla, A. J. Buras and M. E. Lautenbacher, "Weak Decays Beyond Leading Logarithms", Rev. Mod. Phys. **68** (1996) 1125 [arXiv:hep-ph/9512380].

[12] T. Inami and C. S. Lim, "Effects Of Superheavy Quarks And Leptons In Low-Energy Weak Processes $K_L^0 \to \mu\overline{\mu}$, $K^+ \to \pi^+\nu\overline{\nu}$ And $K^0\overline{K}^0$", Prog. Theor. Phys. **65** (1981) 297 [Erratum-ibid. **65** (1981) 1772].

[13] A. J. Buras, "Flavour physics and CP violation", arXiv:hep-ph/0505175.

[14] K. S. Babu and C. F. Kolda, Phys. Rev. Lett. **84**, 228 (2000); S. R. Choudhury and N. Gaur, Phys. Lett. B **451**, 86 (1999); P. H. Chankowski and L. Slawianowska, Phys. Rev. D **63**, 054012 (2001); C. Bobeth *et al.*, Phys. Rev. D **64**, 074014 (2001).

[15] M. Misiak and J. Urban, Phys. Lett. B **451**, 161 (1999); G. Buchalla and A. J. Buras, Nucl. Phys. B **548**, 309 (1999).

[16] M. Blanke *et al.*, JHEP **0705**, 013 (2007); R. A. Diaz, Eur. Phys. J. C **41**, 305 (2005); A. Ilakovac Phys. Rev. D **62**, 036010 (2000).

[17] G. Buras and A. J. Buchalla, Nucl. Phys. B **400**, 225 (1993).

[18] S. L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. D **2**, 1285 (1970).

[19] A. J. Buras *et al.*, Phys. Lett. B **500**, 161 (2001).

[20] C. Bobeth *et al.*, Nucl. Phys. B **726**, 252 (2005).

[21] G. Isidori and A. Retico, JHEP **0111**, 001 (2001); G. Isidori and P. Paradisi, Phys. Lett. B **639**, 499 (2006).

[22] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **94**, 221803 (2005).

[23] Belle Collaboration, M.-C. Chang *et al.*, Phys. Rev. D **68**, 111101(R) (2003).

[24] CDF Collaboration, T. Aaltonen *et al.*, arXiv:hep-ex/0712.1708.

[25] CLEO Collaboration, T. Bergfeld *et al.*, Phys. Rev. D **62**, 091102(R) (2000).

[26] I. I. Bigi, M. A. Shifman and N. Uraltsev, Ann. Rev. Nucl. Part. Sci. **47**, 591 (1997).

[27] I. I. Bigi and N. Uraltsev, Int. J. Mod. Phys. A **16**, 5201 (2001) [hep-ph/0106346]; M. A. Shifman, hep-ph/0009131.

[28] K. Wilson, Phys. Rev. **179**, 1499 (1969); D **3**, 1818 (1971).

[29] W. Zimmermann, Ann. Phys. **77**, 536 and 570 (1973).

[30] A. F. Falk, [hep-ph/9610363], (1996).

[31] A. F. Falk, [hep-ph/0202092], (2002).

[32] N. Isgur, M. B. Wise, Phys. Lett. B **232**, 113 (1989).

[33] J. Chay, H. Georgi and B. Grinstein, Phys. Lett. B **247**, 399 (1990).

[34] T. Mannel and M. Neubert, Phys. Rev. D **50**, 2037 (1994).

[35] M. Neubert, Phys. Rev. D **49**, 3392 and 4623 (1994).

[36] I. I. Bigi, M. A. Shifman, N. G. Uraltsev and A. I. Vainshtein, Int. J. Mod. Phys. **A9**, 2467 (1994);
I. I. Bigi, M. A. Shifman, N. G. Uraltsev and A. I. Vainshtein, Phys. Lett. B **328**, 431 (1994).

[37] A. Ali and I. Pietarinen, Nucl. Phys. B **154**, 519 (1979).

[38] C. H. Jin, W. F. Palmer and E. A. Paschos, Phys. Lett. B **329**, 364 (1994);
A. Bareiss and E. A. Paschos, Nucl. Phys. B **327**, 353 (1989).

[39] A. J. Buras, M. Misiak, [hep-ph/0207131], (2002).

[40] S. Bertolini, F. Borzumati and A. Masiero, Phys. Rev. Lett. **59**, 180 (1987).

[41] N.G. Deshpande *et al.*, Phys. Rev. Lett. **59**, 183 (1987).

[42] M. Misiak, Nucl. Phys. B **393**, 23 (1993) [Erratum-ibid. B **439**, 461 (1993)].

[43] C. Greub, T. Hurth and D. Wyler, Phys. Rev. D **54**, 3350 (1996) [hep-ph/9603404].

[44] K. Adel and Y. Yao, Phys. Rev. D **49**, 4945 (1994) [hep-ph/9308349].

[45] K. Chetyrkin, M. Misiak and M. Munz, Phys. Lett. B **400**, 206 (1997).

[46] C. Greub and T. Hurth, Nucl. Phys. Proc. Suppl. **74**, 247 (1999) [hep-ph/9809468].

[47] A.L. Kagan and M. Neubert, Eur. Phys. J. C **7**, 5 (1999) [hep-ph/9805303].

[48] P. Gambino and M. Misiak, Nucl. Phys. B **611**, 338 (2001) [hep-ph/0104034].

[49] *BABAR* Collaboration, B. Aubert *et al*, Phys. Rev. Lett. **8**8, 101805 (2002); *BABAR* Collaboration, B. Aubert *et al*, hep-ex/0906.2177 (2009).

[50] C. Amsler *et al.* [Particle Data Group], "The Review of Particle Physics", Phys. Lett. B **667**, 1 (2008).

[51] A. L. Kagan and M. Neubert, Phys. Rev. D **58**, 094012 (1998) [hep-ph/9803368].

[52] T. Hurth, E. Lunghi, and Porod, W., Nucl. Phys. B **704**, 56 (2005).

[53] *BABAR* Collaboration, B. Aubert *et al.*, "The BaBar detector", Nucl. Instrum. Meth. A **479** (2002) 1 [arXiv:hep-ex/0105044].

[54] *BABAR* Collaboration, B. Aubert *et al.*, "Observation of the bottomonium ground state in the decay $\Upsilon(3S) \to \gamma\eta_b$", arXiv:0807.1086 [hep-ex].

[55] *BABAR* Collaboration, B. Aubert *et al.*, "Measurement of the $e^+e^- \to b\overline{b}$ cross section between $\sqrt{s} = 10.54$ and 11.20 GeV", arXiv:0809.4120 [hep-ex].

[56] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **99** 021603 (2007).

[57] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **94** 221803 (2005).

[58] Y. Freund and R. E. Schapire, J. Comput. Syst. Sci. **55**, 119 (1997).

[59] R. A. Fisher, Annals Eugen. **7**, 179 (1936).

[60] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. B **241**, 278 (1990).

[61] *BABAR* Collaboration, B. Aubert *et al*, Phys. Rev. Lett. **9**7, 171803 (2006).

[62] *BABAR* Collaboration, B. Aubert *et al*, Phys. Rev. D **77**, 051103 (2008).

[63] Fox, Geoffrey C., and Wolfram, Stephen, "Event shapes in e+ e- annihilation", Nucl. Phys. B **149**, 413 (1979).

[64] *BABAR* Collaboration, B. Aubert *et al*, Phys. Rev. Lett. **101**, 171804 (2008).

[65] A. Ali and C. Greub, Z. Phys. C **49**, 431 (1991); Phys. Lett. B **361**, 146 (1995) [hep-ph/9506374].

[66] A. Ali and E. Pietarinen, Nucl. Phys. B **154**, 519 (1979).